ERIC Document Resume See Instruction Sheet DATE OF RESUME 7/3/65 DATE RECEIVED

A. PUBLICATION DATA

1. ERIC NUMBER ED000863
2. ORIGINATOR CONTROL NO(S). 8408
3. IDENTIFICATION NO. CR 761

DO NOT WRITE IN SPACE BELOW

MICROFILM COPY AVAILABLE YES ☒ NO ☐

COPYRIGHT GRANTED YES ☐ NO ☒

COPYRIGHT RELEASE FILED YES ☐ NO ☐

TYPE OF REPORT
PROPOSAL ☐
IN-PROGRESS ☐
FINAL ☒
FOLLOW-UP ☐

OTHER ERIC REFERENCES:

ERIC SATELLITE CODE

4. AUTHOR(S): Gruber, Howard E. and Weitman, Morris

5. TITLE OF DOCUMENT ABSTRACTED
Self-Directed Study: Experiments in Higher Education
(Cooperative Research Project No. 761)

6. PUBLICATION TITLE: N.A.

7. EDITOR(S): N.A.

8. PLACE OF ORIGIN
ORG. Univ. of Colorado
ST. CITY, STATE Boulder, Colorado
TELE. EXT.
INST/LOC CODE

9. PUBLISHER OR SOURCE University of Colorado

10. VOL.	11. PAGES	12. MO./YR.	13. TOTAL PP.
N.A.	N.A.	Apr. 62	155

B. INDEXING DATA

14. SUBJECT AREA

15. TYPE OF ENDEAVOR experimental study

16. INDEXING TERMS (6-12)

experimental studies G-23	students Y-15
cooperative research R-35	independent study S-14
teaching methods Y-25	

17. ABSTRACT (MAX. 250 WORDS)

A report of classroom experiments conducted over the 3-year period 1957-59 which were designed to study the effects of increasing the college students' intellectual independence and responsibility for their own education. Although the experimental method, "self-directed study," was adapted to meet the needs of courses at different levels and in different subjects, the main feature of the research was the reduction in the amount of time students spent in formal classroom work. Two outstanding features of the experiment were the use of unselected students rather than superior students, and the large number of courses involved in the project. The report describes experimentation involving 19 different courses in 11 different departments.

More than 60 references are listed in the bibliography, and critical comments and suggestions for course improvement appear throughout the report.

ERIC No. 18. SOURCE OF ABSTRACT

C. FUNDING DATA

19. SUPPORTING ORGANIZATION

20. DIVISION OR DEPT.

21. CONTRACT NO.

PROJECT NO.

22. TOTAL COST OF PROJECT

23. DURATION OR PROJECT

U.S. Department of Health,
Education, and Welfare
Office of Education

ERIC Form 1

SELF-DIRECTED STUDY: EXPERIMENTS IN HIGHER EDUCATION

COOPERATIVE RESEARCH PROJECT #761

HOWARD E. GRUBER
MORRIS WEITMAN

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

ED000863

UNIVERSITY OF COLORADO

C. R. P.
761

GRUBER
WEITMAN

REFERENCE
COPY

COOPERATIVE
RESEARCH
PROGRAM

No. 761

UNIV. OF
COLO.

" 761

Self-Directed Study:

Experiments in Higher Education

by

Howard E. Gruber

and

Morris Weitman

University of Colorado Behavior Research Laboratory Report No. 19, April 1962

Self-Directed Study:

Experiments in Higher Education

by

Howard E. Gruber

and

Morris Weitman

University of Colorado Behavior Research Laboratory Report No. 19, April 1962

The research reported in this study was performed in part under a grant from the Fund for the Advancement of Education and in part pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare. Cooperative Research Program. Project No. 761; Principal Investigator, Howard E. Gruber, University of Colorado.

3414A
124

Contents

Preface

Chapter 1 Introduction

Part I Experiments conducted in 1956-1957

Chapter 2 Educational Psychology

Chapter 3 General Psychology

Chapter 4 Civil Engineering: Reinforced Concrete Design

Part II Experiments conducted in 1957-1958

Chapter 5 Freshman English

Chapter 6 Introduction to Research Methods in Psychology

Chapter 7 Physical Optics

Chapter 8 Other Studies Conducted in 1957-1958: Culture and Personality, General Psychology, Independent Study in Honors

Part III Experiments conducted in 1958-1959

Chapter 9 Freshman English

Chapter 10 Physical Science: Man in the Physical World

Chapter 11 Principles of Economics

Chapter 12 Principles of Marketing

Chapter 13 Political Science: American Government

Chapter 14 Principles of Chemical Engineering

Chapter 15 General Psychology - I

Chapter 16 General Psychology - II

Chapter 17 Other Studies Conducted in 1958-1959: Mathematics, Chemical Engineering

Chapter 18 Educational Values, Intelligence, and Self-Directed Study

Chapter 19 Effects of Self-Directed Study on Thinking During Lectures

Part IV The University and Self-Directed Study

Chapter 20 The University as an Environment

Chapter 21 How Students Spend Their Time

Chapter 22 The Goals of Education

Chapter 23 Discussion and Conclusions

References

Appendix A General Scheme for Final Examinations

Appendix B Methodological Notes

Preface

This report presents the results of classroom experiments conducted over the three year period, 1957-1959, together with additional material examining the nature of the institutional setting in which the experiments were carried out. The organizing theme of this research has been the evaluation of the feasibility of increasing the college student's responsibility for his own education by reducing the number of formal classroom meetings. But as the work progressed it became increasingly clear that it was important to pay close attention to clarifying the goals of higher education and to examining the entire set of methods used to further those goals. At the outset we were preoccupied with the assessment of a particular aspect of educational method, but at the end we were more concerned with the network of educational goals implicit in the choice of methods.

We extend our deepest thanks to many agencies and individuals for their support and cooperation. The research was begun under a grant from the Fund for the Advancement of Education; it was continued with the financial support of the United States Office of Education; on several occasions, funds were also contributed by the University of Colorado. Together with former President Ward Darley, almost the entire Department of Psychology played some role in the inception of this study; Professors Marvin Nachman, Kenneth R. Hammond, and Victor Raimy were particularly concerned. Research assistants on the project were William L. King, Stephen W. Link, David S. Gochman, Judith R. Papachristou, and George Stock. In the University administration, Vice Presidents W. Farrell Dyde and Eugene Wilson, Provost Oswald Tippo, and Vice Provost James I. Doi were of great help in a number of ways. We owe a real debt to three Deans for permitting us to invade their academic territories to carry out our experiments: L. D. Coolidge, Dean of the School of Business; Clarence L. Eckel, then Dean of the College of Engineering; and Jacob Van Ek, then Dean of the College of Arts and Sciences. Perhaps our greatest debt is to some thirty faculty members who gave generously of their time and exposed themselves to the kind of scrutiny most teachers are spared; their names are omitted here. We would also like to thank Professors Samuel Baskin and Ruth D. Churchill of Antioch College, and Professor Jerome S. Bruner of Harvard University for their generous help and advice at various stages in our work. Needless to say, none of the above individuals, named or unnamed, is responsible for the research methods we have used or the conclusions we have drawn.

Chapter 1

Introduction

The preeminent system of higher education in the U.S. today involves a very heavy reliance upon methods of teaching in which the teacher exercises close control over the time and thinking of the student. Although the total impact of 4 years of exposure to such methods is not well understood, the general intellectual development of the student may be profoundly influenced by the protracted playing of such a passive role. Ironically, in carrying out his mission in this highly directive manner, the teacher persistently does far more work than his students, rather than serving as a catalyst liberating his students' intellectual energies, and as a guide directing them.

A little simple arithmetic helps to make the point clear. Consider a student attending class 16 hours per week for eight 16-week semesters: this represents a total of 2048 class hours in his college education. Most of this time is spent listening to the teacher talk. Time spent on academic work outside this situation is spent almost entirely in reading assigned textbooks and in other activities which allow little scope for the development of intellectual independence. That this pattern of higher education is not essential is suggested by the fact that the typical British university student, in a system with aims similar to our own, attends less than half as many classes. Far greater reliance is placed on the student's independent studying, both during the academic term and between terms.

The primary aim of the present study is to examine the possibility of increasing the student's responsibility for his own education by reducing the number of formal classes the college student attends. In the present study this reduction was accomplished without altering the formal course structure in which the student's work is divided into five or six courses, each supervised by a different teacher. With few exceptions, the experiments to be reported involve the general run of students to be found at a state university, not a highly selected group of superior students. In preserving this formal structure, and in working with unselected students, the experimental methods to be described differ from those methods that are ordinarily called 'independent study.' In independent study, the student and teacher are in a one:one relationship, with each student pursuing a different course of study or research. Bonthius, Davis, and Drushal (1957) have given a detailed account of the operation of independent study programs in the U.S. In general, the educational result is highly satisfactory, but the method is so expensive that it can be made available to only a limited number of students for a small part of their education. The method used in the present study, which for want of a better name we have called 'self-directed study,' maintains a one:many relationship between teacher and students, thus making possible its extension to a much larger number of students.

In the chapters that follow, studies involving 19 different courses in 11 different departments will be reported. Clearly, no single technique would be applicable under such diverse circumstances. What unifies all these efforts at educational innovation is the attempt in every case to reduce substantially the number of hours the students spend in formal class meetings. As far as possible, we attempted to maintain a standard of reducing class attendance in the experimental group to one-third the number of hours spent by the control groups.

McKeachie (1960, 1962) has recently summarized similar investigations, usually involving single courses in different institutions, but rarely has a similar program, involving many courses, been carried out at a single institution. One important exception is the study conducted at Antioch (Baskin and Churchill, 1961). In many ways, this investigation was similar to the present one, but of course the student body at Antioch and at any large state university can be expected to differ in important respects. A method of teaching appropriate in one setting may not be appropriate in the other. From the Antioch study and from McKeachie's review, the following points emerge: first, under methods of instruction involving sharp reductions in the number of formal class meetings students learn about as much as they do in conventional courses; secondly, self-directed study need not be restricted to exceptionally gifted students; thirdly, self-directed study may sometimes foster certain desirable attitudinal changes. The present study provides an opportunity of verifying these generalizations in a large public university with experiments involving a wide assortment of courses.

In evaluating the effects of self-directed study, account has been taken of changes in the student's knowledge of subject matter, in his ability to think about the subject matter and to learn it independently, and in his general attitudes toward independent intellectual work. The interplay of knowledge and thought has been a frequent concern in investigations of educational methods (see, for example, Thorndike, 1937; Katona, 1940; Kersh, 1958). This issue is taken up specifically in Chapter 7 and Chapters 9 through 16 of the present study. The effect of self-directed study upon intellectual curiosity is a prime concern in Chapters 2 and 3, and, through a fortunate accident, recurs in Chapter 19. In Berlyne's analysis (1954a, 1954b, 1960), curiosity, is treated chiefly as a motivating factor which facilitates learning. But it should be pointed out that the engendering of curiosity need not be considered only as a method of education: it is also a fundamental outcome.

In spite of the enormous weight placed upon the lecture method in higher education, little attention has been given to the student's thought processes in the classroom. The notable and pioneering exception, of course, is the work of Benjamin Bloom (1953). The effects of self-directed study on the stream of thought during a lecture are examined in Chapter 19.

There is widespread agreement that among the most enduring effects of higher education should be changes in the student's attitudes, which in turn may mediate changes in his behavior over periods of time extending far beyond his formal education. Accordingly, the effects of self-directed study on the student's attitude toward doing independent intellectual work have been of constant concern to us and are focussed upon in Chapters 18 and 22. Among the more interesting results to be reported are some surprising relationships between changes in the student's general attitude toward independent intellectual work on the one hand, and his evaluation of self-directed study courses on the other.

Finally, in evaluating the effects of self-directed study, several opportunities were found to examine the persistence of effects of the experimental method. Students who had been involved in self-directed study courses were retested in various ways at intervals ranging from three months to several years after the conclusion of those courses. These results are reported in Chapters 2, 3, 7, 15, and 22.

The effort to evaluate an educational method may be divided into two major parts: first, the accumulation of evidence regarding the effects of the method, such as that described above, and secondly, the decision as to how to use the method. The relation between these two aspects of the process of evaluation is by no means simple. A given method may produce improved educational results, but at a prohibitive cost. A new method may produce a small improvement, and may actually be economical to operate, but the transition from the old to the new may present an imposing obstacle. In the present case, there is no single basis for defining the meaning of positive results. An educator who is convinced that lectures are a highly efficient means of teaching may become more favorably disposed to self-directed study if the difference in effectiveness between one lecture and three lectures per week is vanishingly small. On the other hand, an educator who is searching for new methods of inculcating individual responsibility in the college student may not be satisfied with self-directed study unless the attitudinal changes it induces are substantial and enduring.

In discussions of educational methods, it has become customary to consider the interaction between the characteristics of the individual student and those of the method. Research in this area is summarized in McKeachie's review (1962), with no clear picture as yet emerging. With regard to self-directed study, two important hypothese suggest themselves and will be considered in the chapters that follow: first, the more intelligent student will be better adapted to self-directed study, and secondly, the more self-reliant student will be better adapted to self-directed study. In entertaining these hypotheses, certain cautions should be observed: even if the hypotheses are substantiated, they are not necessarily pertinent for the educational decision-maker. There may be, for instance, no practical way of assigning students with differing degrees of self-reliance to different educational treatments. Moreover, it is not obvious that the student who lacks self-reliance should be protected from situations which may force him to develop it. Analysis of problems of this kind is further complicated by the logical necessity to demonstrate an interaction between a given attribute of the student and a given method of teaching. The demonstration, for example, that a self-reliant student does well in self-directed study is not very revealing if it turns out that he does equally well in conventional courses. If we wish to examine the value of self-directed study for the self-reliant student, we must attempt to ascertain whether or not he does better under that method than he would have done under another method. Similar considerations apply to any other intellectual or personality variables. Results bearing on these problems are presented in Chapters 7 and 18.

It should be a truism that the outcome and meaning of any restricted educational experiment will depend in large part on the institutional setting in which it is conducted. But until very recently, there has been little systematic attention given to the assessment of college environments. New developments in this area have been summarized by McConnell and Heist (1962) and by Stern (1962). Early in the present study an abortive effort was made to use the methods of cultural anthropology to arrive at a description of the institutional setting; this failed, probably for lack of adequate resources to pursue it properly. We then turned to the use of a newly standardized instrument for assessing college environments, the College Characteristics Index. Results obtained with this instrument are described in Chapter 20. Further light is shed on the setting in which the experiments were conducted by the results of an inventory of how students spend their time, reported in Chapter 21. Finally, a study of the goals of education as seen by alumni of the self-directed experiments, reported in Chapter 22, rounds out our description of the university environment into which the experiments were introduced.

Part I

Experiments Conducted in 1956-1957

There must have been something in the air around the Department of Psychology, for in the academic year 1956-57 three of its members, without mutual consultation or full knowledge of each other's efforts, became involved in experiments on the effectiveness of self-directed study. One man conducted an experiment with his own course in Educational Psychology, a second man with his own course in General Psychology. The third, one of the present authors, became involved in a field study of a course in Civil Engineering.

During the same academic year two other members of the same Department began discussions in which almost the whole Department participated leading eventually to the grant from the Fund for the Advancement of Education that supported the continuation of this research in the two following years. Thus it came to pass that the Platzgeist at the University of Colorado merged with the Zeitgeist in a nation-wide effort to study the feasibility of increasing college students' responsibility for their own education.

Chapter 2

Educational Psychology

Spring 1957

Curiosity and Higher Education

Research on college teaching methods has often been concerned with the student's degree of active involvement in the learning process. In such studies, the independent variable is the degree of activity required of the student, as in experiments comparing the lecture method with the discussion method. The dependent variable is usually some measure of the student's mastery of the material of the course. The frequent failure to obtain clear-cut differences in the effects of different treatments may be due to the choice of criterion measures lacking an explicit and well-founded relation to the treatment variable. There is, for example, no obvious reason to expect independent study or other relatively unstructured learning situations to produce greater mastery of a standardized body of knowledge such as is usually tested in final examinations. On the other hand, it does seem reasonable to expect that if greater interest and curiosity about a given field is generated by independent study, it may mediate continuing efforts to solve problems and seek information in that field. In their turn, such continuing efforts may result in superior mastery of subject matter knowledge at intervals remote from the initial learning period. Indeed, this was the result obtained by Kersh (1958) in a laboratory study of the "independent discovery" method.

Thus, curiosity as a directive aspect of cognitive activity assumes special significance in research concerned with long-range effects of educational methods. Unfortunately, curiosity has been given little direct attention in educational research. In some of the older studies cited by Stroud (1940), and in Berlyne's more recent work (1954), the occurrence and timing of questions proposed by the teacher or the textbook has been used as the independent variable, with mastery of subject matter as the dependent variable, but question-raising behavior has never been treated as a dependent variable.

In the present study, the independent variable is the degree of student activity elicited by the method of instruction: self-directed study is compared with a conventional lecture method.

The dependent variables included the usual measures of mastery of course materials. In addition, measures of this kind were repeated after a delay of one year. But the most important departure from previous research was the attempt to measure curiosity as a dependent variable.

Hovey (1958) has conceived of curiosity as arising from the existence of gaps or inconsistencies in cognitive systems. These gaps occur when two cognitive organizations are compared (at various levels of awareness) and are found to be neither wholly congruent nor wholly incongruent.

Several hypotheses concerning the conditions under which gaps and thereby curiosity behaviors would occur were derived from this conceptualization. In brief, the development of cognitive gaps increases as some function of the complexity, number, variety, and relevance-to-self of cognitive systems employed by the individual in dealing with new experiences. The present study is concerned only with the last of these hypotheses. It is hypothesized that psychologically effective gaps are more apt to occur in materials associated with relevant then with irrelevant cognitive systems. By "relevant" is meant more central to the individual's concept of himself, of reality, and of his future. Cognitive systems become relevant to the degree that the individual is required to use them actively in adapting to the situations he confronts. Finally, the intervening variable, cognitive gaps, is coordinated to a behavioral variable: the questions raised by S. Thus, the number and quality of questions generated by S, either spontaneously or in response to E's request to ask questions, was our main behavioral index of curiosity.

An opportunity to test this formulation came to hand as an extension of a self-directed study experiment involving a course in educational psychology. The appropriateness of this situation for testing our hypothesis rests on the assumption that the S in the SDS condition was required more than the S in the lecture condition actively to use and manipulate the course materials, both in his individual study and in interacting with other members of his discussion group. Bloom's (1953) research on students' thought-processes in the classroom lends support to our assumption. He found that during lectures students' thought-processes characteristically take the form of passive recognition and cognizing of the material presented, whereas in discussion groups there is more thinking concerned with problem solution and more comparison of one's own understanding of the material with statements made by others. In addition to measures of question-raising behavior one year after the course, we were interested in the student's direct expressions of interest and in overt behavioral tendencies to seek more information about the field covered in the experimental course.

Experimental Procedure

Subjects. The Ss were two sections of students enrolled in Educational Psychology at the University of Colorado during the spring semester of 1957. Both sections were taught by the same instructor, an assistant professor with five years of postdoctoral teaching experience, including previous teaching of the course now in question. By the flip of a coin it was determined that Section 1, 62 students scheduled to meet at 9:00 a.m. on Monday, Wednesday and Friday, would be taught by a lecture method. Section 2, 53 students scheduled to meet at 11:00 a.m. on Monday, Wednesday and Friday, would be taught instead by the method of self-directed study in small groups. At the University of Colorado, 9 o'clock and 11 o'clock sections are generally considered to be equally desirable, and there is no reason to expect students in these sections to differ systematically.

As entering freshmen, all students had received the University of Colorado Academic Achievement Test (AAT), a placement test giving separate and composite scores in English and mathematics. All students received the

Otis Intelligence Test at the end of the semester in which the experiment was conducted. Near the beginning of that semester, all students received a test of Need for Academic Achievement (NacAch), based on the Liverant test (Liverant, 1958). The lecture group, Section 1, was slightly superior on the AAT and on the Otis, while the SDS group, Section 2, was slightly superior on the NacAch scale. None of these differences was statistically significant.

To test our hypothesis about curiosity, as well as to assess the retention of course material, as many as possible of the original students were administered a questionnaire and a subject-matter test during the spring semester of 1958, about 10 months after the end of the course. Of the original 62 students in the lecture group, 40 were still enrolled at the University, and 36 of these took the post-test. Of the original 53 students in the SDS group, 35 were still enrolled and 30 took the post-test. The groups obtained for the post-test were similar to the groups from which they stemmed with respect to scores on the Otis, the AAT, the NacAch scale, and final course grades.

Treatment. The lecture group met three times a week. Although questions from students were permitted in the lecture class, at least three-fourths of each period was devoted to a planned lecture, which the instructor consciously made more formal than is usual for him.

The SDS class was informed at its first meeting that it had been selected to try out a new method of teaching. Instruction in small self-directed groups was explained and the Ss then received one week of training in the application of techniques which had been found successful in promoting good communication in small groups. Work sheets were distributed which contained concrete suggestions as to the kind of behaviors which were successful in groups and discussion of the project was encouraged in class. In addition, demonstrations of successful group processes were conducted for this class by several graduate students.

After this first week of training in group discussion technique, the SDS class was divided into ten small groups, each having five or six members. Each group was assigned a small room in which to meet each Monday and Wednesday at 11:00 a.m. The groups were equated with respect to AAT and NacAch scores.

The Friday meeting of the SDS class, at which time the instructor met with the entire class, was devoted to a discussion of common problems in group functioning and to questions concerning course content, particularly the latter. Although there was some initial awkwardness, the small discussion groups were meeting smoothly by the end of the second week of the term.

Topics for group discussion were supplied by the course instructor in the form of instruction sheets. These contained a list of the assigned readings, plus supplementary readings for the course. In addition, the instruction sheets contained an outline of the topics to be discussed, and a number of relevant discussion questions. It was suggested that students center their discussions around the instruction sheets. Throughout the semester, new sheets for the next unit of study were distributed to students in the experimental group approximately every two weeks.

When the groups first began meeting, leaders were assigned to each group. These leaders were student members of the group, selected by the course instructor. It was suggested that they keep group attendance records and take the role of leader in group discussions. It was further suggested that the leader's role be rotated throughout the group. In general, groups tended to drop the leadership suggestion after an initial tryout. One or two groups did retain a permanent leader, and most of the groups had the same person fill out the attendance slips from week to week.

Measures of Mastery of Course Material. The final examination consisted of an essay part and an objective part. The course grade was based on the final examination and on tests given during the semester.

The post-test, administered 10 months after the course had ended, assessed retention of course content by means of 15 multiple-choice questions, 24 true-false questions, 10 identification (short answer) questions, and a 10-item matching question. The multiple choice and true-false questions were taken verbatim from the course final and other course examinations. The identification and matching questions were modified versions of questions which had appeared on course examinations. When appointments were made for the post-test, the Ss were given no indication of its nature.

Measures of Curiosity. After the post-test, a questionnaire was administered which included three groups of items related to three aspects of curiosity. The first, and most important, was a group of five items dealing with question-raising behavior. In these five items S was asked to (1) list as many questions as he could think of in the field of educational psychology, (2) comment on a paragraph about an educational problem, (3) list the problems or questions suggested by the previous paragraph, (4) list the things in the field of educational psychology that he would like to find out more about, and (5) list all the questions suggested by another paragraph dealing with an educational problem. Each item yielded a score for the number of questions raised, and the set of five items yielded a score for total number of questions raised. In addition, the protocol was scored for the merit and diversity of the questions raised. During the scoring, the scorer was kept ignorant of the experimental groups to which the protocols belonged.

There were five items in the second group, dealing with interest in the field of educational psychology. Two of these were Likert-type rating scales and three were open-ended questions which were later content-analyzed for expressions of interest in topics related to educational psychology. Again, the scorer was kept ignorant of the groups to which the protocols belonged.

There were five items in the third group, dealing with tendencies toward overt action to get more information about the field of educational psychology. Two of these were self-reports of reading actually done or other actions actually taken, and two dealt with hypothetical efforts the S would be willing to make. The final "action-indicator" was whether or not the S troubled to self-address a postcard asking for information about the outcome of the investigation. These "action-indicators" were used as measures of interest in educational psychology.

Results

Performance in course. At the end of the semester during which the course was taught there was no significant difference between the groups on measures of mastery of course material. The SDS group was slightly but insignificantly superior to the lecture group on the final examination and in the final course grade. A similar result obtained in the post-test, 10 months later. As shown in Table 1, on the post-test, the SDS group was slightly superior in all three parts of the test of retention of course material. None of these differences was significant.

Table 1

Post-test Retention Measures

	Lecture		Self-Directed	
	N=36		N=30	
	Mean	S.D.	Mean	S.D.
Multiple Choice and True-False	37.50	5.06	38.60	4.46
Identification (short answer)	11.38	5.10	12.60	4.97
Matching (Correct/Errors)	1.68	1.16	1.87	1.54

Question-raising Behavior. As shown in Table 2, in each of the five items dealing with question-raising behavior, the SDS group was superior in number of questions raised. None of the differences on individual items was statistically significant. Since the intercorrelations among the scores on these items were negligible, it is reasonable to treat these five scores as independent observations. Using the binomial expansion to calculate the probability of obtaining 5 out of 5 scores all in the predicted direction (Walker and Lev, 1953) the superiority of the SDS group is significant ($p < .03$). In addition, the SDS group was superior in the diversity of questions raised, in the total number of questions raised, and in the merit of the questions raised.

Table 2

Number of Questions Raised

	Lecture		Self-Directed	
	N=36		N=30	
	Mean	S.D.	Mean	S.D.
Item				
1. "What questions do you have about the field of psychology as applied to education? Put down as many questions as you can think of."	3.03	1.90	3.53	1.30

(Table 2 continued)

Item	Lecture N=36 Mean	Lecture N=36 S.D.	Self-Directed N=30 Mean	Self-Directed N=30 S.D.
2. "Read the following selection and then comment on it in at least 25 words." (Selection concerned problem of child in second grade.)	.53	1.29	.73	1.48
3. "What problems or questions did the selection you just read (in Item 2) suggest to you?"	3.42	1.74	3.90	1.18
4. "Is there anything in the field of Educational Psychology that you want to find out more about? What?"	.75	.80	1.03	.90
7. "Read the following selection and then in the space below write as many questions as you can think of that the quotation suggests to you." (Quotation asserted need to emphasize mathematics and other "basic" subjects in the high school curriculum.)	2.75	2.11	2.86	2.25
Sum of Questions Raised, Items 1, 2, 3, 4 and 7.	10.58	4.12	12.10	3.87
Merit of Questions (Sum of judgments of merit of response to items 1, 2, 3, 4 and 7).	15.88	2.33	17.16	2.91

Table 3 shows the frequencies of individuals raising different numbers of questions, subdivided into low, middle, and high numbers of questions raised. Although the middle categories are of about the same size in both SDS and lecture groups, the proportion of individuals in the extreme categories is in accord with the hypothesis that the SDS condition stimulated greater curiosity as evidenced by question-raising behavior.

Interest

In four of the five items dealing with interest in the field of educational psychology, the SDS group was superior, and in four of the five items dealing with tendencies toward overt action indicating interest in educational psychology the SDS group was superior. Thus, the SDS group was superior in 13 of the 15 items in the post-test questionnaire. Since the intercorrelation among scores in the various items were negligible, it is legitimate to treat these 15 scores as independent observations. Using the binomial expansion, this result is highly significant (p .004).

Table 3

Students Classified by Total Number of Questions Raised

Total Number of Questions	SDS	Lecture
Low (3-8)	4	12
Middle (9-14)	17	19
High (15-23)	9	5

Chi-square = 4.76

$.05 < p < .10$

Action-indicators

Within the group of "action indicator" items an interesting trend emerged. The SDS group reported a larger mean number of readings since the course than did the control group, while the control group reported a larger number of activities other than reading. This difference in pattern was highly significant (chi-square = 10.95, $p < .001$). The readings reported included serious books, professional journal articles, articles in serious magazines (e.g., Harpers), and articles in slick magazines and newspapers. Other activities reported included observations of children in and out of school, observations of teachers and other people, application of course material to the teaching of piano and remedial reading, attending school board meetings, and attending lectures. Most of the activities other than reading were concerned with the application of what had been learned. Thus, the significantly different distribution of activities reported suggests that SDS students, in the year following the course, had tended to do more things to increase their knowledge, while the control students had tended more often to apply what they had learned.

Discussion

Discussion of this experiment is deferred until after the next experiment is reported, since the two are closely related.

Chapter 3

General Psychology

Spring 1957

Curiosity and Higher Education Continued

Since the results of the experiment above were promisingly consistent but of only tenuous statistical significance, it was decided to repeat the study with another course. By a happy accident, during the Spring of 1957 another instructor had taught two sections of the second semester of General Psychology, one section by a conventional method, one section utilizing self-directed study. As in the above experiment, the SDS section met only once a week, while the control section met three times a week. The SDS section was organized in small discussion groups, usually of 5 or 6 students, which met at a place and time of their own choosing, with no instructor present.

Because the decision to repeat the previous study, focussing attention on curiosity as the dependent variable, was taken after the course in General Psychology had already been given, arrangements could not be made to obtain comparable final examination scores from both groups. Instead, the groups were matched on the basis of the Academic Achievement Test (AAT) given to all Freshmen and the results to be reported deal only with the post-test, given one year after the course was finished. The students in the post-test consist of 14 matched pairs; in every case, either both members have identical scores on the AAT, or where identical scores could not be located the pair was constructed so that the SDS member would have a slightly lower score, thus working against the hypothesis that the SDS group would exhibit superior performance in the post-test.

The post-test consisted of an objective test of retention of material covered in the second semester of General Psychology; items intended to elicit question-raising behavior; items dealing with interest in psychology; and items dealing with action-tendencies toward getting more information about psychology. Although all of these kinds of items have been described above, the specific items were modified to suit the changed material of the course. For example, in one of the question-raising items, the Ss were asked to comment on a passage describing the "authoritarian personality," rather than on a passage about educational psychology. Finally, the post-test was somewhat shorter than in the previous study, to reduce labor.

Results

On the retention test, the SDS group was slightly but insignificantly superior to the control group.

On each of three items dealing with question-raising behavior, the SDS group was superior in the number of questions raised. In the total number of questions raised, the superiority of the SDS group just failed of significance, using a two-tailed Wilcoxon test for paired replicates. The SDS mean for the total number of questions raised was 8.0, the control mean was 6.4.

On each of three items dealing with expressions of interest in psychology, the SDS group was superior to the control group. On one of these the superiority of the SDS group was significant at the one per cent level, using Wilcoxon's test for paired replicates. This item was a nine-point rating scale for the question, "In general, how interested are you in Psychology?" The midpoint of the scale was 5, the SDS mean = 5.41, the control mean = 4.18. On the other two expressions of interest the difference was not statistically significant, although it was in the same direction.

There were three action-indicators, on two of which the SDS group was superior to the control group and on one of which the reverse was true. None of these differences was significant.

Thus, on 8 out of 9 items, the SDS group was superior to the control group in interest and curiosity about general psychology.

Discussion of Chapters 2 and 3

It seems clear that self-directed study produced a small but persistent superiority in question-raising behavior and in other indices of curiosity. In both studies when tested almost a year later, the SDS groups asked more questions and were prepared to exert greater efforts to seek information about educational psychology. Although they had only one-third the usual number of contact hours with the instructor, the SDS groups suffered no loss in retention of factual knowledge.

While the differences produced by the experimental treatment are rather small, it should be remembered that these were differences that survived for almost a year even though they result from a relatively restricted change in the student's educational program. The fact that these results parallel Kersh's findings in a laboratory setting (1958) and resemble those found recently by Beach (1961) in another classroom experiment lend added weight to the notion that self-directed study fosters curiosity.

It is worth remarking that the instructors in the two SDS sections were quite different. There was a ten year age difference between them. The older instructor would unquestionably be described as mature, steady, and well-organized--a man for whom the lecture might seem a natural medium. The younger instructor would be described as youthful, ebullient, and enthusiastic--a man who particularly enjoys contact with individual students. In spite of these differences, essentially the same results were obtained in both courses.

The question remains, when should we expect greater curiosity to lead to more enduring mastery of subject matter? Curiosity cannot lead to information-seeking and continued mastery of material if the environment presents S with demands that are antagonistic to independent information-seeking. Thus, isolated experiments in self-directed study can have only limited effects. But if curiosity is not the only attribute of the permanently educated individual it is at least a *sine qua non*. In this sense, curiosity should not be viewed as an end in itself but as an essential class of events in the stream of cognitive processes that characterize the intellectually independent individual.

Surely a major goal of education is to stimulate the student to further pursuit of knowledge on his own initiative, after the compulsions of the classroom are far behind. Curiosity may be said to have a "gatekeeper" function in the educational system: if the system arouses further exploratory or re-organizing behaviors, it may set off a process which is self-sustaining and which may in large part determine the whole character and direction of the individual's future life.

Chapter 4

Civil Engineering: Reinforced Concrete Design

Spring 1957

In discussions of engineering education--whether with engineering students or faculty--one frequently encounters the opinion that, especially in this field, students must receive precise, up-to-date information and that this requirement cannot be satisfied by teaching methods that permit the student considerable freedom and latitude, such as discussion groups, projects, and independent reading. Paradoxically, however, professional engineers are keenly aware of the explosive growth of technical knowledge and the consequent impossibility of teaching the student the specific things he will need to know when he leaves the university and takes up his first job. Moreover, part of the intellectual equipment of a university-trained engineer, one would hope, is the confidence and resolution necessary to continue his own education in order to take advantage of the growth of knowledge. Self-directed study, therefore, seems more imperative and more dangerous in engineering education than in other fields. No wonder one student, on first hearing of the experiment described below, called it "frightening"!

Experimental Procedure

The conventional course. This experiment concerns a one-semester elective course for seniors in the College of Engineering. Ordinarily, the instructor spent the first half of the semester on the basic theory of reinforced concrete design, using lectures and short, assigned projects as the major teaching techniques. The students received fairly close supervision and guidance. Usually, the second half of the semester was spent on covering certain special applications of the basic theory, the teaching techniques remaining about the same. About once a month, three-hour tests were given during a laboratory period, these tests stressing the solution of problems. The formal organization of the course entailed three lectures per week and two three-hour laboratory periods per week, but the instructor used some of the laboratory time, possible as much as half, for lecturing, generally in connection with the short problems the students were working on.

Self-directed study. In the semester of the experiment, the instructor, after a discussion with one of the present authors, decided to give the students an opportunity to spend a significant portion of their time in self-directed study. After the first half of the course (basic theory) had been completed in the ordinary way, the students were given four weeks to work on a single problem, to be chosen by them from a group of six problems formulated by the instructor. The problems were intended to simulate full-scale design problems, about the same as might be encountered in professional work. It was expected that the students would have to dig out a lot of new material on their own. The students were asked to work in pairs. Regular lectures and laboratory meetings were suspended for this period, except for occasional progress reports. The instructor remained available for consultation at regular class times.

The instructor in the course was a man in his early thirties with a master's degree in civil engineering and several years of previous experience, both in teaching and in practical work. At the time of the experiment, he was finishing work on his doctorate, and has since spent a year on a postdoctoral fellowship. It would be fair to describe him as a man with an unusually high academic aspiration level, both for himself and his students. The students quite uniformly respect him for his enthusiasm as well as for his knowledge and ability; they also consider him to be a hard man who stresses detail and gives them too much work to do.

There was no control group available. Nevertheless, it was felt that questionnaires administered to the students just before and just after the self-directed study project would shed valuable light on the nature of independent study as applied to engineering courses. For administrative reasons, later efforts to repeat the experimental procedure with an adequate control group were unsuccessful.

The instructor inaugurated the project with an introductory talk which was characteristically enthusiastic and seems to have effectively caught the students' interest in participating in the new departure. The class was then turned over to a psychologist who administered an open-ended questionnaire. He began by explaining the research value of the questionnaire and of the project as a whole, and attempted to dissociate the evaluation of the experiment from the instructor and from the grades the students would receive in the course. Reassurances were given that their questionnaire responses would be kept confidential and not revealed to the instructor; honest, expressive answers were called for; the content of the responses suggests that such answers were given.

Results

Initial questionnaire. The first questionnaire was administered just after the plan for the self-directed study project had been explained by the instructor. It centered on two points: the students' reaction to the course up to that point and their reactions to the new departure that had just been explained. In view of the exploratory nature of this study, only four open-ended questions were asked, and the results to be reported were obtained by a very simple analysis of the written responses.

On the whole, the students' reaction to the course was very positive. Out of 23 respondents, 13 indicated that they were enjoying the course, 8 were ambivalent, and only 2 were not enjoying it. On the question of the relevance of the course to the student's professional goals, 15 felt it was definitely relevant to their goals, 2 felt it was somewhat relevant, and 6 felt it was not relevant to their goals. Good points most frequently mentioned generally centered around the instructor's ability and his clear and enthusiastic coverage of material, both theoretical and applied. Bad points most frequently mentioned were concerned with time-consuming assignments, unnecessary detail, and the instructor's attempt to cram in as much content as possible. In short, the students felt the course was worthwhile, the instructor very good but very demanding.

The students' expectations concerning the self-directed study project were quite favorable, 17 indicating a definitely favorable attitude, 5 an ambivalent attitude, and only 1 an unfavorable attitude. From the comments, it would appear that the instructor, in introducing the project to the students, had spoken well to their condition. They saw the project as a relief from the heavy supervision characteristic of other courses, as an opportunity to try their hands at a project simulating professional problems they might soon encounter, and as a chance to test their own abilities and commitments in the field of engineering. The following are some typical comments indicating these feelings:

Subject 4: "I am delighted with this proposal. It should help me to find out if I could enjoy design work, by dealing with a project rather than disjointed pieces. At this point I feel that a drawback to the system is that I will be completely ignorant of certain types of design."

Subject 7: "It seems that this may cause a lot of work but may stimulate and get a lot out of the person. But also I may not feel this way if it turns out to be too much, causing my other courses to fall down."

Subject 9: "I like the idea of this type of project. It helps to find one's own abilities and more nearly parallels actual practice."

Subject 15: "The project seems to be a step in the right direction and I hope it will shed some light on motivation problems of my own."

Subject 17: "At first glance I think I like the idea. You certainly will have more time to go into a 'thing' more thoroughly and really understand it better. This is superior to many little problems, I think."

The misgivings expressed by Subjects 4 and 7 were common. Engineering students feel that they are hard-pressed, and they fear innovations that may increase their work load or leave them ignorant of things supposedly vital to their future work.

Final questionnaire. The final questionnaire was administered when the students reconvened for the third phase of the course, a few final lectures on special topics. Their impressions of the value of the project to them is of special interest because at the time the questionnaire was given, the students' projects had not been returned to them, and they did not know the grades they had received.

Of 24 students responding, 16 indicated that they had enjoyed the project, 3 were ambivalent, and 5 had not enjoyed it. The amount of time reportedly spent on the project ranged from 17 to 85 hours, with a mean of 40 hours. The students were asked to compare the amount of time they had spent on the project with the time they would have spent on the course during a comparable period in the first half of the semester. Of 22 students responding, 14 indicated that they had spent more time on the project, and 8 indicated that they had spent the same or less time. The time it took was the most frequent complaint levied against the project, but there was no apparent relationship between expressions of this complaint and the student's enjoyment of the project.

In general, the division of work between the members of the pair seems to have been equitable, with only one student reporting that he had done less than 40 per cent of the total work and no one reporting that he had more than 67 percent of it. The students were also asked to indicate how they had divided their time between "background research and theory" and "numerical work and drawings." All the students indicated that they had done some of both, the most extreme division of time reported being 80 percent of time on the former category, 20 percent on the latter. This student remarked, "Since I chose a topic untouched in class work a good deal of reading was necessary." Sources of information consulted included textbooks, reference works, pamphlets, handbooks, engineering codes, periodicals, the instructor in the course, other instructors, outside experts, and other students.

In the second questionnaire, the students were asked to "compare the benefits derived from this project with the benefits derived from a comparable period of the first half of this course." The following are the comments of students representing the entire spectrum of opinions fairly accurately.

Subject 4: "More topics were covered in a comparable period of formal class work, with a lot less work and a little better understanding. However, the knowledge gained in class work only lasts as long as the first test. The subject I studied is firmly implanted in a way which can only be done by having to dig something out for yourself."

Subject 7: "Compared to the first half I learned very little, if anything. On parts of the subject we had there was not too much written, and I believe I would have gotten more out of this last half if we had had lectures and examples. I believe the problems given were too broad and in some cases too large to get accomplished in the length of time required since there were some other courses that had to be finished at the same time."

Subject 9: "Perhaps the only benefits gained from this project were confidence and originality. However, with the excellent lecturer we have for this class the time was wasted in terms of knowledge and understanding. With a poorer lecturer the time may have been well spent on the project."

Subject 10: "I don't believe the knowledge derived from the last half would compare to that of the first half. It is harder to answer questions that arise than it would be if they could be asked and answered orally. ...It has increased confidence, however, simply because I know that I can get the needed information by myself."

Subject 15: "(a) Knowledge - hard to compare since we applied to these problems what we had learned during the first half. No comparison. (b) Understanding - more understanding gained through use of basic principles on some 'concrete' design. (c) We compared our design with an actual design for this particular tank and ours seems just as good if not better. Wonderful feeling of confidence and assurance and idea of our abilities on a commercial job. (d) The project tied together a lot of loose ends and connected the various courses together into one force which enabled us to work the problem (diagram illustrating summation of forces). (e) I'd say that once the basic essentials are learned, this independent study is far superior to the ordinary accepted methods."

Subject 21: "The benefits of this project are far greater than those of the conventional class. We may not have learned as much, but it gave me a chance to do what we may do later on. Working with a partner has helped me to do the work. I was then responsible not only to myself but to my partner."

The group was about equally divided in their estimates of the relative merits of the two systems. Of 22 students responding, 10 felt they had gotten more out of the "new system," 3 thought they were about equal, and 9 felt they would have gotten more out of the old system. But even among those who felt they would have learned more in the conventional system, there were some who felt that self-directed study should be retained, because it had distinctive values not attainable in their other courses. The students were asked, "Do you advise retaining and possibly enlarging this type of approach? Could you suggest any improvements?" Of 23 students responding, 15 thought self-directed study should be retained, 6 thought it should not be, and 3 gave ambiguous answers. The comments and suggestions for improvements are illuminating. The following are a fair sample:

Subject 4: "Yes, continue this approach. Limit the problems as to their scope. In practice the problem I did might well have taken an office of designers months to complete. Therefore the job I handed in was incomplete, which detracts from the feeling of accomplishment."

Subject 9: "I would advise retaining this type of approach with improvements. One improvement would be to make the projects more precise and direct. Another improvement may be to make the problems suitable for one person."

Subject 13: "Personally I do not advise this type of program. If it were possible to forget entirely the other courses which one takes, for a period of say three weeks, and spend this time entirely on such a project, the idea might be okay."

Subject 15: "I definitely am a proponent of this approach after basic essentials are learned. One big improvement would be to cut down the extent of some problems. Our particular problem was just about the right 'size'--maybe even a little too small--but some of the other guys had too much to do for the time allotted."

Subject 17: "I think it would be hard to retain one course on this basis where the rest were on the old system. In the old system there is a requirement of regular and successive assignments and time must be regularly scheduled for them. To try to work one course like this into the system would be difficult, because, I think, to accomplish very much long periods of uninterrupted study are necessary. To try to accomplish a large project with accompanying research would be seemingly prohibitive if you can only spare an hour now and then."

Subject 21: "I don't advise enlarging the method in this course. It should be continued and possibly used in other courses."

Almost every student had suggestions for improving the experimental method, touching upon: choice of project, availability of reference materials, time of introduction of project, working alone or in groups, number of projects, availability of the instructor for conferences, continuation of some lecturing during the project period, and extension of the method to other courses. Considering the number of suggestions the students found to make, it is all the more impressive that such a high proportion of them felt the method should be retained.

Personal relevance and response to self-directed study. The small number of cases and the non-quantitative character of the data do not admit of any very extensive efforts to find correlations between the subjects' responses to different questions. But a few variables were tabulated against each other, with suggestive results. As might be expected, those who enjoyed the project a great deal tended to recommend the retention of the experimental method more unreservedly than those who only enjoyed it moderately. The modal response is moderate enjoyment coupled with the opinion that the method should be retained with modifications. Of the three individuals who actually disliked the project, two thought the method should be discontinued. Using Fisher's exact method, the relationship between enjoyment of the project and opinion of SDS is highly significant ($p < .01$). These results, shown in Table 1, suggest that enjoyment of the project was a necessary but not a sufficient condition for a favorable opinion of self-directed study.

Table 1

Enjoyment of Project and Opinion of Self-Directed Study

	Enjoyment			
	High	Moderate	Neutral	Negative
Opinion of SDS				
Unqualified approval	5	1		
Qualified approval or doubtful	2	9	1	1
Disapproval	1	2		2

It will be remembered that on the first questionnaire, the students were asked to indicate whether or not the course was relevant to their own professional goals. Table 2 shows the relation between the degree of relevance and the student's reaction to the experimental method. On the whole, those for whom the course was highly relevant took a firmer stand for or against the experimental method than did those for whom the course was not so relevant. Combining the two extremes, this tendency is significant beyond the 5 per cent level, using Fisher's exact method. Thus, among the students for whom the course was highly relevant, the experimental method often produced a strong reaction, but not always a favorable one. Some felt that it spoke to their condition, others felt that it deprived them of an opportunity to hear a particularly good lecturer on topics that would be important to them.

Table 2

Effect of Relevance of Course to Professonal Goals on

Opinion of Self-Directed Study

	Relevance	
	High	Not High
Opinion of SDS		
Unqualified approval	5	1
Qualified approval or doubtful	4	8
Disapproval	4	1

Instructor's reaction. The instructor's view of the course is not unlike the students'. He approached the experiment with some trepidation, fearing that the students might react to the period of self-directed study by slackening their efforts in his course, or that they would be unable to work independently. In spite of these fears, he was extremely pleased with the students' projects, especially when it turned out that some students who had wholly disappeared from sight had done far more ambitious and creditable work than had been expected of them. With minor modifications, the instructor continued to use the experimental method in subsequent semesters. Later, however, the course was reduced from five semester-hours to four, and the project was eliminated. Instead, a new course was developed, Advanced Structural Design, the chief purpose of which is to give students an opportunity to do design projects. This new course was a direct outgrowth of the present study.

Discussion

It can be concluded that the experiment with the independent study project in this course was, on the whole, rather successful. The following points should be noted: 1. The course is a senior engineering course attracting a fairly homogeneous group many of whom will soon be taking jobs and consider the course highly relevant to their professional goals. 2. The independent study project had some of the appearance of a very practical project, approximating professional working conditions more closely than other academic work. 3. The instructor is highly respected by the students and was able to "sell" the project to them in a very effective introductory talk.

One recurrent complaint was the fear that the time consumed in the project meant that certain specific topics vital to professional work were being slighted. Had the students actually understood the impossibility of acquiring in advance the specific knowledge necessary for specific jobs, they might have seen the function of a university education in a different light.

Part II

Experiments Conducted in 1957-1958

During the academic year 1956-57 an inderdepartmental group at the University drew up a proposal for University-wide research on self-directed study. Departments involved in the original proposal were: Economics, Honors, Philosophy, Physics, and Psychology. When plans matured and research funds were obtained, a faculty seminar was organized to discuss the proposed research. In addition to the departments mentioned above, faculty members joined the seminar from the following departments: Anthropology, Biology, Chemistry, Civil Engineering, English, History, Mathematics, Political Science, and Sociology.

The meetings consisted of informal discussions centering on the goals of higher education, capabilities of students for independent work, and the extent to which faculty members should permit their teaching methods to be influenced by economic considerations such as the increase in college enrollments and the consequent shortage of teachers. There was considerable feeling that training students to be intellectually independent might be a good thing in its own right but that the faculty should not permit itself to confuse this worthy educational goal with the need to solve the teacher shortage. There was much discussion, but no agreement, on whether reduction in the number of lectures would reduce the teacher's expenditure of time or would lead to a costly increase in conferences with individual students. There was fairly general acceptance of the hypothesis that the brighter students might adapt well to self-directed study, but the less intelligence students might suffer from it. The term, "self-directed study" was generally disliked; the faculty felt that even if there were a sharp reduction in classroom work, the teacher should still have an important directive function. The faculty were interested in "increasing the student's responsibility" rather than in allowing himself to direct his own studies.

It was apparent that not all those participating in these discussions were interested in educational experimentation or in reducing the number of hours spent in the classroom: rather than supporting these goals of the project, many faculty members simply welcomed the opportunity to discuss problems of higher education. As one remarked, speaking for many, this was the first time in his career as a teacher that he had been part of a group interested in discussing the process of teaching.

In the course of the next two years, the various departments involved in the original research proposal participated in the actual program to varying degress. One department that had committed itself to a large-scale experiment with a sophomore course found, when the final decision had to be made, that they could not agree to reducing the number of lectures in a sophomore course. Instead, they substituted a course generally taken by juniors and seniors. Another department, although addicted to much informal experimentation, never gave a course that could be considered within the compass of this study. It may be that the members of this department were really in fundamental disagreement with the aims of the study; certainly, several of them voiced the view that the teacher's presence was essential for the kind of discussion that was the heart of learning in that field. On the other hand, departments joined the project that had not been involved in it at the beginning.

The plan for 1957-58 was to conduct as many controlled classroom experiments as possible using parallel sections taught by experimental and conventional methods. The single criterion for a group designated as "experimental" or "self-directed study" (SDS) was to be a marked reduction in the number of hours spent in the classroom. The research group actually had in mind a reduction from approximately three hours per week to one hour per week in class. As matters turned out, in 1957-58 there were three experiments involving parallel sections: in Physical Optics one group attended almost no classes, the other attending three times a week; in Freshman English there was a reduction in class time, but not as much as hoped for, although student responsibility was increased in other ways; in Introduction to Research Methods in Psychology both groups attended the same three lectures per week, but there was a sharp difference in their laboratory experiences. In addition to these courses involving parallel sections, other studies were conducted involving two courses in Psychology and Independent Study for students in the Honors Program. All of these will be described in the chapters that follow.

Early in the Fall semester the students involved in these various studies were invited to participate in a testing program connected with the educational research program. To minimize the "Hawthorne effect" the letter of invitation was framed in such a way as to convey the feeling to the control subjects that they too were signled out for special attention; there is, of course, no need to belabor the point, which will be discussed later, that the SDS subjects knew quite well that they were participants in an experiment. Our problem was to minimize the important of this factor by giving the control students the same impression insofar as possible.

Tests administered in the Fall included: the Army General Classification Test (Civilian Form), the Edwards Personal Preference Schedule, the Dressel Test of Critical Thinking, and a questionnaire concerning certain background information. Early in the Spring semester of 1957 the same students were invited to another testing session at which were administered an alternate form of the Dressel Test and the STEP test of reading comprehension. Considering the voluntary nature of this testing program the respondent rate was quite high, 205 out of 252 potential respondents completing the whole battery.

As the reader will discover, much of the effort expended on collecting the above data was wasted, simply because most of the classroom studies conducted in 1957-58 were not well enough controlled to warrant analysis of the effects on performance of subtle variables of personality and attitudes. The one course where such an analysis was undertaken, with some interesting results to be described below, was Physical Optics, both because this experiment was better controlled and because it was more extreme than any of the others.

How the vicissitudes of the Self-Directed Study Project in 1957-58 shaped its course the following year will be described in Part III. The chapters now following take up, one by one, the classroom studies conducted in 1957-58.

Chapter 5

Freshman English

Fall 1957

> "a girl asked abruptly, as if disaster threatened, 'Are we going to have grammar?'" Hughes Mearns in Creative Power: The Education of Youth in the Creative Arts.

When Monsieur Jourdain suddenly discovered that he had been speaking prose all his life, he may have been impressed, but the rest of the world was not. In the same vein, educators have sometimes wondered whether a seat of higher learning was an appropriate place to repeat instruction in the essentials of the mother tongue that the student has previously encountered by precept in elementary and secondary school, and by example from dozens of teachers. But the question may be pressed even further: is it, perhaps, not only inappropriate but useless--even harmful--to drill college students in the rules of grammar?

Experimental Procedure

English Language 100. All entering freshmen at the University of Colorado take an English language achievement test on the basis of which they are assigned to different levels of instruction. Those receiving an unsatisfactory score on this test are assigned to a remedial course. Those doing exceptionally well on the test may enroll for an advanced course emphasizing the creative and critical use of language skills. Most students, however, are assigned at random to sections of English Language 100, a course emphasizing the acquisition and practice of basic language skills. The objective of the course is to bring the student up to or above a certain minimum standard of correct and effective use of the English language. In the year of this experiment, the course was intended to give the student practice and instruction in writing themes, in making speeches, and in grammar.

Self-directed study. Viewed as an innovation, this experiment was a relatively modest departure from previous practice. The students in the experimental group were given very great freedom of choice of topics for compositions (themes); small study groups were formed whose function it was to criticize their members' prepared speeches before presentation to the class; all formal drill in grammar in the classroom was eliminated; occasionally, about once every two or three weeks, class meetings were cancelled to allow additional time for library research and study group meetings. The control groups met regularly three times a week.

It was originally planned that the same teacher should teach the experimental section and a control section. But she soon reported that it was not entirely possible for her to maintain a clear distinction between the teaching methods used in the two classes. In particular, the control group was not receiving as much class instruction in grammar as was usually the case; in general, the students in the experimental group were reacting so well to the additional responsibility they were being given that she found it hard to withhold the same teaching methods from the control group. Consequently, another section which was being taught in the conventional way by a comparable instructor (also a woman, and also having more than 15 years of teaching experience) was enlisted to serve as a control group. This teacher was known to put considerable emphasis on formal training in grammar. The study therefore involves three groups: a "no grammar" experimental group and a "some grammar" control group taught by the same instructor, and a "much grammar" control group taught by another instructor. In the order named, the three groups received increasing amounts of formal class instruction in grammar.

Results. All students in English Language 100 took the same final examination, one part of which was devoted to composition and one part to grammar. Arrangements had been made to have all papers of the students in the three sections involved in the experiment graded by one person who would remain in ignorance of the students' identities while grading. Unfortunately, these arrangements fell through, and each teacher graded her own students on the composition, thus making meaningful comparisons impossible for that part of the final.

In scoring the grammar test the intended precautions were taken. Table 1 shows the results. The first column gives the mean scores of the three groups on the AGCT, a test of general intelligence given early in the semester. It can be seen that the three groups are quite comparable. On the grammar test, the "no grammar" group was best, the "some grammar" group was next best, and the "much grammar" group was poorest. This analysis of variance, shown in Table 2, indicates a significant difference among groups. The difference between the "no grammar" group and the "much grammar" group is highly significant ($t = 3.09$, $p < .01$); the other differences are not significant. In short, the more class time spent on grammar, the more poorly the class did on a test of grammar at the end of the semester.

Table 1

Performance of English Language Sections on Various Tests

Group		AGCT		Reading Comprehension		Grammar Final	
	N	M	SD	M	SD	M	SD
Experimental: no grammar	23	121.8	9.6	26.7	3.8	84.3	5.3
Control A: some grammar	24	122.7	8.0	26.4	3.9	82.0	7.3
Control B: much grammar	23	122.6	5.8	26.5	3.6	77.4	9.0

There was no appreciable difference in the students' satisfaction with the three sections.

In an interview after the semester, the teacher of the experimental seminar expressed the opinion that most students have had years of instruction in grammar

Table 2

Analysis of Variance of Performance on Grammar Final

Source	df	MS	F
Between groups	2	287.0	5.02*
Within groups	67	57.2	
Total	69		

$^*p < .01$

and are able to continue with a minimum of guidance. On the other hand, few freshmen have had much opportunity to analyze, organize, and express ideas. In her view, the experimental group had not received enough guidance and practice in the development of these skills, and she planned to give more of such instruction in the future. She regarded the evidence as conclusive that formal classroom drill in grammar was a waste of time at the college level.

Discussion. Are these results plausible? Are they to be trusted? Do they link up with anything else in our experience? Perhaps. In the far-reaching "Eight-Year Study," one high school English department developed an experimental program in which the students participated in a voluntary reading program. Required readings, formal book reports, formal training in grammar, and similar devices were eliminated. These were replaced by a host of livelier and more informal procedures to foster the students' interest in reading and writing. A day each week was "free reading day." With whatever guidance they requested, students read books of their own choice. Classroom libraries were developed and the school library improved. Over a three year period, students in the experimental group read three times as many books as did those in the control group. The teachers concluded, "Never again will students in the Altoona High School think of reading as forced labor" (Aikin, 1943, p. 8). Of especial interest are the results of tests of the students' mastery of grammar. The experimental group, which had received no formal training in grammar, did markedly better than a carefully matched control group which received three years of "very definite training in formal grammar with college entrance examinations in view" (Aikin, 1943, p. 10).

In a later chapter we will present some material on the stream of thought in the classroom. Our evidence and that of others suggests that the teacher has very much less than complete control over the minds of the students in his classroom. The amount of mind wandering is sometimes great; classes in which the students are engaged in active discussion produce attentive listening and self-critical inner thoughts, as contrasted with the more passive responses of students supposedly listening to a lecture; students who attend fewer lectures are more attentive and thoughtful listeners than students who attend many.

Several lines of evidence suggest that repetitious practice sometimes inhibits the performance being practiced. For example, one of Lewin's students, Karstens (1928), performed experimental work on "psychic satiation", showing that repetitious work carried on ad nauseum produces profound intellectual as well as emotional disruptions. Such laboratory experiments concentrate the satiating experience in an hour or two, but there is at least room to believe that the same cognitive disruptions may be produced by educational experiences less concentrated but more protracted.

Finally, the whole tenor of modern psychological thought, ranging from Freud to Lewin, suggests that a student who feels that a topic or a form of instruction is puerile, or beneath him in other ways, may withdraw psychologically from the instruction that is offered him. Bodily withdrawal from college by able and successful students is not rare (see, for example, Gruber, 1961), but our studies of students' thinking in the classroom suggest that mental withdrawal is an everyday affair.

The enthusiastic and justifiably exasperated teacher of the language skills may well argue that the typical freshman has not earned his right to withdraw from the study of an essential skill until he has mastered it. Nevertheless, if it is true that certain methods of instruction provoke in the student a complex chain of withdrawal responses, other methods must be found.

Chapter 6

Introduction to Research Methods in Psychology

Fall 1957

Experiments conducted with a single course suffer from the defect that the students are involved in a number of other courses not under the investigator's control, which may have unknown or unexpected effects on the experiment. This problem is especially prominent in the present experiment which illustrates well the difficulties of conducting studies of this kind, but has some interesting results nonetheless.

Experimental Procedure

Introduction to Research Methods. This is a course taken by Sophomores and Juniors beginning a major in psychology. A year of introductory psychology is taken previously or concomitantly. The course covers experimental design, statistical methods, and other aspects of research methods. The major goals of the course, as expressed by the instructor involved in the experiment were: first, to make the student reasonably conversant with the way scientists work and what they are trying to accomplish, with the nature of the problem of scientific induction, and with some of the basic techniques employed in scientific psychology; secondly, to make it possible for the student to read and comprehend the literature of scientific psychology, including standard psychological journals. As the course had been taught previously, the students attended two lectures per week and one three-hour laboratory period. Instruction in statistics was heavily emphasized in the lectures. The laboratory period was devoted to repetitions of standard psychological experiments, generally at the rate of one experiment per week, although two weeks were allotted for some experiments. Occasionally, laboratory time was used for extra instruction in statistics.

Self-Directed Study. For purposes of the experiment, the 38 undergraduates enrolled in the course were divided into 19 pairs, matched on age, sex, year in school, major, grade-point average, and on a test of knowledge of scientific method. This test was a fictitious experiment with numerous flaws in it, the task being to discover the flaws. One member of each pair was assigned at random to the experimental group, the other to the control group thus forming two laboratory sections. The control group received the treatment outlined above.

The experimental group attended the same lectures as the control group. For the first five weeks, both groups had the same laboratory work. After that, the experimental group was dismissed from all further laboratory sessions. Instead, their task was to design two original experiments and conduct one of them, including a written report comparable in form to a published journal article. Both the lecturer, an assistant professor, and the laboratory instructor, a graduate assistant, were available for consultation.

The best laid plans of mice and men ... One of the students in the experimental group died. Unknown to the investigators until the middle of the semester, two of the students in the experimental group and eight of the control group were taking a laboratory course in the Psychology of Learning. As it happened, the spirit in which that course was conducted closely resembled the aims of self-directed study. To illustrate, the students in the course on learning were given a mimeographed sheet which read, in part, as follows:

"Our intent in the laboratory work is to help students develop the skills for research on problems in learning. These skills include the discovery and formulation of hypotheses, design of research plans, search of the literature, utilization of available equipment, collection of data, analysis of data, and writing research reports.

"...After the first few weeks there will be no more planned laboratory sessions. The instructor and assistant will have regular hours set aside, however, to help you with all phases of your (original) research throughout the semester."

Quite clearly, almost half of the control group was getting a type of experience similar to that intended for the experimental group.

To add to the confusion, the morale of the control group sagged very seriously around the middle of the semester, due in large part to their feeling that the laboratory was "just busy-work" (the Assistant's phrase). To ameliorate the situation, they were given one very open-ended assignment, to design an experiment. According to the Assistant, many members of the control group were upset by the sudden change to independent work. But on the whole, morale improved, and for the rest of the semester the students in the control group did brief experiments planned by their teachers.

The net result of these efforts to conduct an experiment were approximately as follows: the control group received somewhat more formal classroom instruction in topics in statistics and experimental method, since some of their laboratory time was devoted to this purpose. The control group conducted a series of 8 instructor-designed experiments, following detailed instructions prepared for them. The experimental group all had at least one major opportunity to conduct an independent piece of research, with very little guidance. Some members of each group conducted an independent piece of research in the course on Learning. If they were in the experimental group of the Research Methods course, this meant they did two independent researches. If they were in the control group of the Research Methods course but also in the Learning course, they did one independent research. Those students who were in the control group but not in Learning had no opportunity to do independent research.

Results

The two groups gave almost identical performances on the final examination, which stressed topics in statistics and experimental methods. There was no direct comparison possible between the laboratory performance of the two groups, since their tasks had been so different. The Assistant felt that the control group had learned more of standard information about experimental method, but as the semester progressed he became increasingly unsure of the worth of this knowledge as compared with the experience of working independently.

The students in both groups were dissatisfied with the course. Some of this feeling probably stemmed from the instructor's superior manner in the lectures; when he became aware of this problem and attempted to change his behavior, morale improved. Some of the dissatisfaction probably stemmed from the existence of the experiment, members of each group wanting some of what the other group had. Students in the control group objected to the lack of freedom to exercise creativity and independence. Students in the experimental group objected to the lack of structure and guidance. The instructor and Assistant agreed that both criticisms were valid.

One aftermath of the course is worth considering. A continuation of the Research Methods course was offered the following semester. Of the 37 students involved in the experiment, 18 enrolled for the second semester. If we consider the course in Learning to have been a form of self-directed study as well as the experimental group in the Research Methods course, there is a greater tendency for students who had some self-directed study in one form or the other to enroll for the additional course than for students who had had no self-directed study. Combining the first three columns of Table 1, 15 out of 26,or 58 per cent of the students who had had some self-directed study enrolled for the second course; in the control group, those who had had no self-directed study, 3 out of 11, or 27 per cent enrolled for the second semester. Using Fisher's exact method, this is not quite significant (p = .087).

Table 1

Student Continuation in Second Semester of Research Methods Course as Influenced by First Semester Treatment

	First Semester Treatment			
Behavior	SDS and Learning	Control and Learning	SDS only	Control only
Continued	2	5	8	3
Did not continue	-	3	8	8

Discussion

The final examination indicated no difference between experimental treatments. The examination covered the standard topics, in which the control group had received somewhat more classroom instruction than the experimental group. Unlike our later experiments, no specific provision was made for measuring gains that might have reflected the experimental group's greater opportunity for independent work and thought. For the present, it is merely worth noting that two groups receiving different amounts of instruction performed equally well.

Students enrolled in the Research Methods course who had some self-directed study were more likely to enroll for the second semester of the course than students who had none. Taken by itself, the weight of evidence on this point is not very convincing: with the number of students involved, the difference between 27 and 58 per cent is not statistically significant. But taken together with results presented earlier, the evidence appears to mount that exposure to self-directed study does produce some increase in curiosity about the subject studied.

Possibly the most important result of this experiment was the change in educational philosophy that took place in the instructor and Assistant. At the outset, they were opposed to the spirit of self-directed study, believing that students need formal classroom instruction on all important points, coupled with specific assignments, tests, and all the other paraphernalia of teacher-directed education. They were willing to participate in the experiment, but simply to be cooperative in a colleague's research efforts, not out of support for self-directed study. This attitude had two important consequences. In the control group the students were over-supervised to the point of crisis, although possibly this crisis would not have become apparent to the teachers if they had not been involved in the experiment. The crisis led them, at one point, to introduce into the control laboratory a method of instruction that might have been appropriate as an introduction to self-directed study. Meanwhile, the experimental group, rather than receiving adequate preliminary instruction and guidance, were treated on a sink-or-swim basis, leading to serious morale problems in that group as well. The difficulties produced by this treatment in the experimental group were handled by giving some of the students extra time to finish their work, extending the allotted time into the following semester.

In coping with these difficulties, the instructor and Assistant changed in three important ways. First, they became much more favorable to self-directed study and more relaxed at the prospect of allowing students to work independently. Secondly, they realized that they had gone to extremes in their sink-or-swim approach, and that some guidance was necessary in helping the student to work independently. Finally, they became more alert to the need to match teaching techniques with objectives. For example, it became apparent that the goal of teaching students to read the literature could be better pursued by having them do so early in the semester, rather than by a surfeit of textbook assignments about the literature.

In immediately subsequent semesters, not included in this study, the plan followed by the instructor included continuation of the basic approach used with the experimental group but with some additional guidance and occasional laboratory meetings throughout the semester. These modifications appear to have produced more satisfactory results.

Probably the most valuable lesson to be drawn from the present study is that college students are ill-prepared for intellectual independence, and that the teacher fostering self-directed study must be prepared to cope with complaints of the kind the students made in this course. Transitional techniques are needed to smooth the change from conventional courses to more independent work. Moreover, independent work in an educational setting should not be construed to mean total withdrawal of the teacher's guidance.

Chapter 7

Physical Optics

Fall 1957

A common finding in educational research is the absence of any difference in performance between groups treated differently. One explanation of this result is the use of tests designed to measure changes in conventional subject matter knowledge although the experimental treatment may have been designed to foster progress along other lines. Another explanation might be the cautious character of many experiments: the experimental treatment does not really differ very much from the control. The present study avoids both of these difficulties. The final examination was designed to test knowledge and skills at four points along a continuum ranging from specific factual knowledge at one extreme to the ability to learn independently and apply new material at the other, on the hypothesis that the former favors the conventional group, while the latter favors the self-directed study group. The conventional group met regularly, three times a week, while the experimental group had no formal classwork at all after the first two weeks. So long as we remain within the framework of organized courses, there is not much room for a more radical departure from the usual program of formal class meetings.

Experimental Procedure

Control group. The control group received a very standard course covering the first half of a standard textbook in Physical Optics. The students received three lectures per week, two one-hour tests during the semester, and assignments of problems to be done at home. Students were called on in class and graded on the way in which they did the homework problems at the blackboard. In an interview after the course, the instructor estimated that "about 60% of the students did the homework faithfully, and 40% only occasionally." Sometimes, if a question arose in class, a student (usually the one asking the question) was assigned to look up the subject and report on it in class. This method, the instructor feels, is especially useful with historical topics. The instructor made a special effort to keep his lectures up to a consistently high level, "for fear," as he put it "that if I let down on this part of the work it wouldn't be a good experiment." As compared with his usual lectures, "they were no worse, and probably better because I tried a little harder to keep the level up." Although we have no other information about the quality of these lectures, it is probably safe to assume that they were better than average for the university as a whole. The instructor is a full professor, a well-known and vigorous research worker, and much in demand for lectures around the campus and elsewhere.

Self-directed study. After the first week of the semester, the experimental group did not hear any lectures on optics by the instructor. They were given a set of five major topics indicating the material that would be covered in the course; these topics were not identical with chapter headings in the textbook and therefore obliged the student to search through the textbook and other books for the appropriate material. Students in the experimental group were not required to use any particular book, but in general they used the text assigned to the control group, supplementing it with other references, some of which they got from the instructor.

The experimental group was divided into study groups of four or five members each, the groups being approximately equal in composition. Each group elected a chairman. Near the beginning of the semester, the instructor met with the group chairmen twice; in these meetings each chairman presented his group's plan for studying optics. After that, there were no further group meetings with the instructor until about three weeks before the end of the semester, when he had a one-hour meeting with each group. In his words, the time was spent "going over the topics, pointing out to them some of the aspects of the physics involved that they certainly should be familiar with, asking them if they were familiar with it, and finding that in most cases they did have some idea of the topic but were not quite sure of the intricate details that I went over in class." This was to give them some indication of the depth of understanding that might be called for on an examination. At the end of the term, there was one meeting of all students in both experimental and control groups, where students were free to ask any questions at all. Not many questions were asked, except about the type of final examination to expect.

Throughout the semester, the instructor and assistant were available for consultation and questioning, a privilege not often exercised. Some of the study groups stopped meeting regularly after the first month or so, their members feeling that it was more profitable to work alone, but they tended to resume meetings toward the end of the semester. Other groups met regularly throughout.

Assignment of subjects. At the end of the first week, all the students enrolled in the course were tested on the first chapter of the textbook--in other words, on their ability to study optics independently. On the basis of this test, together with their grade point average and major subject (whether physics or engineering), two matched groups were formed, one designated as self-directed study, the other as control.

After the experiment was under way, two members of the control group and two members of the experimental group dropped out of the course, and one student requested transfer from the experimental to the control group and then dropped the course. None of these students took the final examination. Lastly, one student transferred from the experimental to the control group and did complete the course; it was decided to eliminate him from consideration in the experiment as he was the one individual whose membership in a group was based on his own decision. The net result, then, was an experimental group of 22 Ss and a control group of 28 Ss.

Final examination. In discussions with the instructor, it was decided to construct the final examination to test performance at four points along a continuum of complexity of thought. These points were, in the instructor's words:

"Part I. Simple facts and ideas on optics.

Part II. Applications of above facts and ideas to solve somewhat simple problems.

Part III. More difficult applications.

Part IV. Entirely new material - exposure, then questions."

Both groups received the same final examination.

Results: Group Differences

Morale. There was nothing unusual about the attitude of the Control group toward the course. In the Experimental group, as mentioned above, two students dropped out early in the semester, and two asked to be transferred to the control group. The instructor granted their requests. One of these students was in military service and indicated that he simply could not respond to a self-directed study situation. There was one student whose father, a scientist, fearful that his son's grades would suffer, brought pressure on the Physics Department to have the son transferred to the Control group, but the boy remained in the Experimental group (and got a B in the course).

In general, however, the members of the Experimental group accepted the fact that they would receive no lectures and would have to work out the subject for themselves. Although they were somewhat at a loss as to how to proceed, they managed to formulate plans and get on with the work. After about four weeks, when they discovered that independent work was difficult, and that they were not going to be able to coast through the material, the dropouts and transfers began to take place. Several student contacts with the research group, as well as the comments of the instructor and assistant reflected the occurrence of a drop in morale at that time. After that, the morale rose and the students in the SDS group remained confident that they could do the work on their own. In the instructor's words, when he met with the study groups toward the end of the term, "I got the feeling that they had everything under control and were not afraid of the final exam."

On the negative side, some of the students objected to the self-directed study procedure. In the assistant's words, "They felt that it was unfair for them to be paying a certain amount of tuition and not getting a certain amount of instruction for that tuition. And also they thought it was unfair that they were to be graded on the course and yet not have any instruction." From some of the comments, there is reason to believe that these complaints were intensified by the fact that some students were arbitrarily assigned to the self-directed study treatment while others, studying the same subject at the same time, received a more conventional course. In any case, apart from the students mentioned above, most of the students accepted the experimental procedure and went ahead with it.

Final examination. The students' performance on the final examination is shown in Table 1. It can be seen that the Control group was slightly superior on Parts I and II while the Experimental group was superior on Parts III and IV. Only the difference on Part II was statistically significant.

In order to test the significance of the difference in the pattern of performance of the two groups, a single score was computed for each student by subtracting the sum of the scores on Parts I and II of the final examination from the sum of

Table 1

AGCT and Fall Semester Final Examination Scores of Optics Class

			Final Examination				
		AGCT	I Facts	II Simple problems	III Complex problems	IV Learning new material	Exam. Total
Self-directed study	N	20	22	22	22	22	22
	M	136.3	68.9	43.6	60.9	63.0	236.4
	SD	6.4	11.0	19.4	19.9	18.1	46.1
Control	N	27	27	27	27	27	27
	M	135.1	74.4	54.0	56.7	57.0	242.0
	SD	9.5	10.9	13.8	19.1	22.5	51.6
	Diff.	1.2	5.5	10.4	4.2	6.0	5.6
	t	.52	1.67	2.06	.75	1.01	.39
	p	n.s.	n.s.	$< .05$	n.s.	n.s.	n.s.

the scores on Parts III and IV. In the resulting score, negative numbers then represent a pattern in which the student did better work on the simpler parts of the test, while positive numbers represent a pattern in which better work was done on the more complex parts of the test. These scores were than arranged in a contingency table as shown in Table 2. The difference in pattern is highly significant, the

Table 2

Distribution of Patterns of Performance on Final Examination in Optics

(Parts III + IV) - (Parts I + II)

	≤ Md	> Md
Self-directed study	6	16
Control	20	7

Chi-square = 8.86

$p < .01$

Control subjects having done their best work on the simpler material while the Experimental subjects did their best work on the more complex material.

Follow-up in Spring Semester. Some of the students in the Optics course in the Fall semester enrolled for a continuation of the course in the Spring. One of the main functions of the Spring semester course was to give the student laboratory experience, in addition to which certain topics in optics were covered in greater depth than in the Fall. Students originating in the Fall semester experimental and control groups were thrown together as one class in the Spring. The instructor in this course, who had become interested in the self-directed study experiment, gave the students a surprise examination at about the middle of the semester. Part of this test corresponded to Part I of the Fall semester final examination, both as to the form and content of the items. Part of the test corresponded to Part IV. The final examination had three parts, corresponding to Parts I, II, and III of the Fall semester final.

Table 3 shows the results on these tests, together with the AGCT scores for the students involved. It can be seen that the groups were still quite well-matched in intelligence. On the surprise test those students who had been in the Experimental group displayed the same pattern of performance as before: they did their best work on the more complex material. Those who had been in the Control group also persisted in the previously displayed pattern: they did their best work on the simpler material.

Table 3

AGCT and Spring Semester Performance of Students

Continuing Optics Course

			Surprise Test		Final Examination		
Fall semester treatment		AGCT	Facts	New material	Facts	Simple Problems	Complex Problems
Self-directed study	N	11	12	12	12	12	12
	M	135.6	74.5	60.0	44.8	12.8	10.9
	SD	8.1	10.9	28.6	5.9	4.2	4.6
Control	N	14	14	14	14	14	14
	M	134.7	80.1	48.2	45.0	15.1	11.1
	SD	11.2	14.1	12.9	12.3	4.5	4.5
	Diff.		5.6	11.8	.17	2.3	.16
	t		1.10	1.25	.04	1.29	.09
	p		n.s.	n.s.	n.s.	n.s.	n.s.

To test the significance of the difference in pattern of performance on the mid-term examination, a difference score was computed for each student by subtracting the score on the "Facts" test from the score on the "Learning New Material" test. The difference between the two distributions, Control and Self-Directed Study groups, was statistically significant, as tested by the Mann-Whitney U Test (U = 37, $p < .02$).

By the time of the final examination in the Spring, the differences between the two groups have almost disappeared. Their performances are almost identical on questions corresponding to Parts I and III of the Fall semester final. Only on Part II is there any appreciable difference, the Control group remaining slightly superior on simple problems, as they were in the Fall semester, although the difference is no longer statistically significant. Unfortunately, this final examination was not designed in consultation with the research group; had it been, it might have been possible to include a section involving learning new material.

Results: Individual Differences

In this section, the number of cases varies slightly because not all students took all tests. Since this section deals with individual differences rather than group differences, the one student who transferred from the Experimental to the Control group and completed the course, who was eliminated from the analysis in the preceding section, is now included.

Intelligence and Performance in Course. Scores on the AGCT ranged from 112 to 150, the second lowest score being 122. This narrow range is to be expected in a group that has survived in college for three or four years. In the following pages, when we refer to students of "low" or "middle" intelligence, the reader should keep in mind that we are referring only to position within a rather homogeneous sample. In spite of the small number of cases involved, we first inspected the data with the students divided into three groups: low, middle and high intelligence. Previous experience in other studies (Gruber, Brady and Means, 1959; Gruber, 1960) had shown the importance of examining educational data for the possibility of curvilinear relations between intelligence and other variables.

Table 4 shows the relationships between AGCT and total final examination performance. It can be seen that there is no obvious, simple relationship. In the Control group, there is a very slight tendency for the most intelligent to do better than the middle or least intelligent; in the Experimental group, however, it is the least intelligent who do best, while the middle intelligent seem to suffer by comparison with the others. None of these trends is statistically significant, but further analysis of the sub-parts of the final examination is suggested.

Table 4

Intelligence and Total Final Examination Performance in Optics

Contingency Table

	Final Exam Performance	112-129	130-140	141-150	Total
Control Group	> Median	4	5	6	15
	< Median	5	5	3	13

(Table 4 continued on next page)

Table 4 (continued)

	Final Exam Performance	AGCT 112-129	130-140	141-150	Total
Group	> Median	4	3	2	9
	< Median	1	8	2	11

following presentation, we have grouped Parts I and II as the parts of the final examination representing simpler performances, and Parts III and IV as the parts representing more complex cognitive performances. In each pair, the sum of the two scores was used.

Table 5 shows the relationship between AGCT and performance on Parts I and II of the final examination. Here the pattern is similar to that obtained in the total final examination score, but again, none of the trends is statistically significant.

Table 5

Intelligence and Performance on Parts I and II of

Final Examination in Optics

Contingency Table

	Final Exam Performance	AGCT 112-129	130-140	141-150	Total
Control Group	> Median	5	6	6	17
	< Median	4	4	3	11
Experimental Group	> Median	3	2	2	7
	< Median	2	9	2	13

Table 6 shows the relationship between AGCT and performance on Parts III and IV of the final examination. The striking fact about the table is the success of the low intelligence students in the Experimental group, both as compared with their counterparts in the Control group and as compared with the more intelligent members of the Experimental group.

Since the division of the cases into three levels of intelligence reduces the number of cases in some cells too drastically for statistically significant

Table 6

Intelligence and Performance on Parts III and IV of Final Examination in Optics

Contingency Table

	Final Exam Performance	AGCT 112-129	130-140	141-150	Total
Control Group	≥ Median	4	4	6	14
	< Median	5	6	3	14
Experimental Group	≥ Median	5	5	2	12
	< Median	0	6	2	8

relationships to appear, the data in Table 6 have been regrouped, as shown in Table 7. Table 7a deals only with cases below the median on AGCT; Table 7b deals only with cases above the median on AGCT; cases exactly at the median have been omitted.

Table 7

Interaction of Treatment, Intelligence, and Performance on Parts III and IV of Final Examination

Contingency Tables

a. Students Below Median on AGCT

		Experimental	Control
Final Examination III + IV	> Median	6	3
	≤ Median	2	12

b. Students Above Median on AGCT

		Experimental	Control
Final Examination III + IV	> Median	4	8
	≤ Median	5	4

It can be seen that students below the median on AGCT did well on Parts III and IV if they were in the Experimental group and badly if they were in the Control group. Using Fisher's exact test, this pattern is statistically significant ($p < .025$). Among students above the median on AGCT, the opposite pattern appears, but not significantly: the more intelligent students tended to fare better in the Control group than in the Experimental group.

There are two commonly encountered expectations regarding the relation between intelligence and performance. First, it is expected that the more intelligent students will do better than the less intelligent students. Secondly, it is expected that the more intelligent students will respond more favorably to opportunities for independent work. As far as the present results go, neither of these expectations is borne out. In the Control group, there is possibly a very slight, non-significant tendency for the most intelligent Ss to do better on Parts III and IV of the final, but on Parts I and II all levels of intelligence performed alike. In the Experimental group, the least intelligent Ss do best on Parts III and IV. On Parts I and II the middle intelligent Ss suffer as compared with the other groups.

Probably the most important result of this section is shown in Table 7. Students below the median on AGCT who could do quite well on Parts III and IV (requiring complex thought and independent learning) if assigned to the Experimental group, do very badly if assigned to the Control group.

Level of Aspiration and Performance in Course. In a questionnaire administered near the beginning of the semester, the students were asked two questions which may be considered as indicating their "academic level of aspiration." One question asked: "Realistically, what grade point average are you trying to attain in the courses you are taking this semester?" The other asked "What grade point average do you think the average student in your classes is trying to attain in the courses he is taking this semester?" Since both responses may be considered as reflecting the student's own aspiration level, the sum of the two responses is used as the measure of aspiration level in this section. The median value for Experimental and Control groups combined was almost exactly 6, reflecting the fact that the students in this course are striving for grades of "B" and assume that other students are doing likewise. There was no difference between Experimental and Control groups in aspiration level, nor was there any appreciable correlation between AGCT scores and aspiration level.

For Experimental and Control groups combined, there is a statistically significant tendency for students with a high aspiration level to do better work. As shown in Tables 8 and 9, this is true for Parts I and II and for Parts III and IV of the final examination.

There seems to be a complex interaction between treatment, aspiration level, AGCT, and performance. As far as these data go, they suggest that in the Experimental group, a combination of high aspiration level, with low AGCT was most likely to lead to good performance on Parts III and IV, while in the Control group, a combination of high aspiration level with high AGCT was most favorable. Although some of these relationships are statistically significant, we are omitting the numerical results and tests of significance, because the number of cases involved in such detailed breakdowns is too small to permit us to ignore the possibility that these trends were only random fluctuations appearing in a set of data inspected from many angles. Nevertheless, the trends referred to suggest an interesting hypothesis which will be taken up in the discussion.

Table 8

Aspiration Level and Performance on Final Examination Parts I and II in Optics: Contingency Table for Experimental and Control Groups Combined

		Aspiration Level	
		$<$ Median	$>$ Median
Final I + II	$>$ Median	7	16
	$<$ Median	15	9

chi-square = 6.22

$p < .02$

Table 9

Aspiration Level and Performance on Final Examination Parts III and IV in Optics: Contingency Table for Experimental and Control Groups Combined

		Aspiration Level	
		$<$ Median	$>$ Median
Final III + IV	$>$ Median	7	15
	$<$ Median	15	10

chi-square = 4.95

$p < .05$

Personality and Performance in Course. Near the beginning of the semester, the students enrolled in this course took the Edwards Personal Preference Schedule (Edwards, 1959). This test (EPPS) provides measures of 15 personality variables originating in Murray's list of manifest needs (Murray, 1938). Recently, Milton and Lipetz (1961) have factor analyzed the EPPS and shown that the list of 15 needs can be reduced to 5 factors. In developing a composite score for each of these factors, we computed a simple sum of the individual's scores on the three variables most heavily loaded on a factor, excluding those variables which did not load in approximately the same way for men and women. The list of factors as named by Milton and Lipetz, together with the variables we used in the composite score for each factor, are presented below. A minus sign preceding a variable merely indicates that its sign is reversed in computing the individual's composite score.

Name of factor	EPPS Variables composing factor
Interpersonal relationships and affiliation	Succorance, Nurturance, Heterosexuality
Aggression	Aggression, -Nurturance, -Affiliation
Status-dominance	Dominance, Achievement, Exhibition
Structure and orderliness	Order, Endurance, Deference
Freedom and independence	Autonomy, Chance, Intraception

As shown in Table 10, the Experimental and Control groups are quite similar to each other in their composite scores on the 5 factors, none of the differences approaching statistical significance. It is evident that the students in this course differ appreciably from Edwards' normative sample of 760 college men. For the reader's convenience, Table 10 includes those norms computed for the 5 factors, based on data given by Edwards (1959). The most impressive differences are the greater need for structure and orderliness, the greater need for aggression, and the lesser need for interpersonal relationships expressed by the students in the experiment than by the normative sample.

Table 10

Strength of Needs Measured by Edwards Personal Preference Schedule:

Experimental and Control Groups Compared With College Norms

Factor[1]	Experimental N = 23		Control N = 29		College Males[2] N = 760
	M	S.D.	M	S.D.	M
Interpersonal relations	37.39	8.30	36.67	6.05	42.43
Aggression	47.47	9.86	49.00	7.97	43.75
Status-dominance	47.04	7.26	49.27	9.20	47.50
Orderliness	41.17	13.85	39.68	9.28	34.10
Freedom and Independence	47.60	7.29	47.51	7.49	45.97

[1] Factors are groups of needs based on factor analysis of EPPS.

[2] Computed from norms published by Edwards (1959)

The needs expressed by the students were examined for the possibility of a relationship of personal need with performance on the final examination. This was done first by graphical inspection and by a series of median tests. No statistically significant trends emerged. The only trend approaching significance was a tendency for students in the Control group expressing a low need for freedom and independence to do better on Parts I and II of the final than students expressing a high need. Since the relationship suggested is quite important, the matter was pursued further by a two-way analysis of variance in which the two independent variables were scores on the factor, freedom and independence, and the educational treatment. There was no significant effect of the freedom and independence factor, nor, what is more important, was there any significant interaction between the factor and educational treatment.

In spite of these negative findings, the possibility of an interaction between need and treatment upon performance was pursued still further by re-analyzing the component need scores that had gone into the freedom and independence factor.

For two of these needs, *autonomy* and *intraception*, no particular relationships could be discerned. For the third need, *change*, an interesting trend appeared. Table 11 is a contingency table summarizing the results for the relation between need for change and the final examination. The most striking feature of both tables is the fact that the students having a high need for change do relatively poorly on the final examination. This is especially true of performance on Parts I and II for students having a high need for change who were assigned to the Control group.

Table 11

Interaction of Treatment and Need for Change upon

Performance on Final Examination

Contingency Tables

		Experimental Group				Control Group		
		Need for Change				Need for Change		
		Low	Mid	High		Low	Mid	High
a. Performance on Final Parts I and II	High	2	3	2	High	6	9	1
	Low	3	5	5	Low	2	1	7
		Need for Change				Need for Change		
		Low	Mid	High		Low	Mid	High
b. Performance on Final Parts III and IV	High	3	6	2	High	5	3	3
	Low	2	1	5	Low	2	6	6

Various efforts were made to clarify the picture by considering patterns of needs, such as need for achievement in relation to need for freedom and independence. None of these efforts bore much fruit beyond the plethora of hypotheses always produced when problems of this richness remain unresolved.

To sum up, there is some not very strong evidence of an interaction between need for change and treatment in their effects on performance; there is much stronger evidence for an effect of that need on performance in both groups. It is probably unnecessary to warn the reader that there is no very satisfying way of assessing the statistical significance of results eventuating from the inspection of so many variables. The reliability of these results could only be tested by cross-validation in a subsequent experiment, but it would be no mean feat to replicate the essentials of this particular study. The trends reported here, therefore, should be taken as merely suggestive of possible lines of further inquiry. In general, this effort to make sense of the relation between personality and performance in an educational experiment yielded very meager, essentially negative results.

Summary and Discussion

In a course in Physical Optics one group of students (Control) attended classes three times a week and another group (SDS) attended no classes after the first week, being required to learn the material for themselves. The final examination included sub-tests at four points along a continuum ranging from specific factual knowledge at one extreme to the ability to learn independently and apply new material at the other. On the final examination the Control group did better on tests of factual knowledge and solving simple problems while the SDS group did better in solving complex problems and in learning and applying new material about optics. This difference in patterns of performance persisted for at least two months, as evidenced in an examination given to students who were enrolled in a continuation of the optics course.

Contrary to a widely held hypothesis, students getting lower scores on a test of intellectual ability seemed to profit the most from self-directed study. This was especially true of performance on parts of the final examination involving complex thought and independent learning: low scorers on the AGCT did well if assigned to the self-directed study group, badly if assigned to the control group. It should be noted that "low" scorers within this group were still rather able individuals, since the mean AGCT score for the entire group of students was 136.

To obtain a measure of academic level of aspiration, students were asked to indicate the grade point average they were trying to attain and also the average that the average student in the same classes is trying to attain: these two estimates were summed. Combining SDS and Control groups, students with high aspiration level gave superior performances on the final examination. In the SDS group, a combination of high aspiration level and low AGCT was most favorable to good performance on the final examination; in the Control group the most favorable combination was high aspiration level and high AGCT.

In an analysis of the relation between personality and education an attempt was made to verify the hypothesis that students appearing to be more self-reliant on a test of personality would adapt better to self-directed study, while students appearing to be less self-reliant would adapt better to more conventional methods. The hypothesis was not confirmed. In an extended analysis of the relation between performance in the course and performance on the Edwards Personal Preference Schedule, the only personal need that seemed to be related to course performance was the need for change: students having a high need for change did relatively poorly on the final examination; this was especially true of students in the Control group for the parts of the final examination involving factual knowledge and simple problems.

The need for change is described as the need "To do new and different things, to travel, to meet new people, to experience novelty and change in daily routine, to experiment and try new things, to eat in new and different places, to try new and different jobs, to move about the country and live in different places, to participate in new fads and fashions" (Edwards, 1959, P. 11). The present results suggest that the students high in this need, probably most restless and jaded with years of education in an unchanging, perhaps stultifying mold, can no longer fulfill their own intellectual promise. This is especially true when they are given routine tasks and more than enough help from the teacher.

The results of this experiment suggest that the student with a high level of academic aspiration may develop poor study habits if he discovers that he gets enough help from his teachers to do well without studying hard. This is more likely to happen to the brighter students than to those who are not quite so bright. For the very bright student, regular class attendance in a conventional course, together with desultory study habits and "cramming" for the final examination, is a successful adaptation to the university environment. If the environment is suddenly changed in a way that puts a premium on greater individual responsibility, the old adaptation is no longer successful. On the other hand, the student who, because he is not such a facile learner, has had to develop a more responsible pattern of studying, can more successfully make the change to an environment demanding greater independence. In our experiment, the transition from the more protective to the more demanding environment is effected by the contrast between the usual run of courses and the self-directed study course. This transition resembles the shift from the university to the world outside. Successful adaptation to a highly protective environment does not necessarily augur well for later adaptations to a more complex world. By requiring the student to live through the confusion and initial disorientation of self-directed study, we may foster the development of mature patterns of work that are more suitable after he leaves the university.

Chapter 8

Other Studies Conducted in 1957-58

Culture & Personality, General Psychology, Independent Study in Honors

In the three preceding studies parallel experimental and control groups were available, and reasonable quantitative comparisons were possible. In three additional studies to be presented in this chapter no such comparisons are possible. The studies will be described briefly, partly for completeness and partly because they do add something to our total understanding of problems involved in increasing the student's responsibility for his own education.

Culture and Personality (Psychology)

This is a two-credit course intended for students not majoring in psychology (psychology majors usually take Social Psychology instead). Enrollment is restricted to Juniors and Seniors. Previous background in psychology ranged from 0-25 credit hours.

The initial intent of the instructor and the research group was to explore the optimal amount of guidance that should be given by the teacher to students engaged in small group projects.

The class was scheduled to meet twice a week. Instead, the instructor asked them to form groups of 3-5 students for the purpose of studying together and to investigate an aspect of the course in which they were interested. Each group was required to present a report of their investigation to the rest of the class. With 60 students enrolled, 16 groups were formed. The class met for only one lecture per week, the time saved being used for the group meetings.

When the groups were formed, the instructor--unknown to the students--designated half of the groups as "guided" groups, the other half as "unguided". He met with the guided groups very early in the semester to help them plan their work, suggesting bibliographies and methods of work to them. All groups met with the instructor shortly before they gave their reports, but in his meetings with the unguided groups the instructor simply answered their questions, while in his meetings with the guided groups he played a more directive role. Thus, to some groups the instructor appeared to direct their work, to make sure that they made group decisions and knew where they were going; to other groups, the teacher was merely available for consultation.

As a pretext for making appointments with some groups and not with others, the instructor scheduled the guided groups to give their reports first.

In other respects, the course was hardly an example of self-directed study. A dittoed sheet was handed out, giving week-by-week reading assignments in two textbooks. The instructor characterized his own lectures as "such as to discourage interruptions." Three tests were given during the semester, as well as the final examination.

This disparity between the self-directed study generated by the group projects and the teacher-directed character of the remainder of the course reflects the explicit point of view of the instructor. In an interview he stated that "one of the goals of education as a whole is development of an attitude in an individual that he can learn about problems in which he becomes interested, and he can think about practical problems from a theoretical and empirical point of view." The projects were intended to foster progress in that regard. But lectures and readings planned by the teacher have a function too: "I feel that if students are encouraged to follow their own interests there is an extremely advantageous result, namely that these interests mature and broaden and become autonomous. But there can be a disadvantage in such self-directed study if this is the only content of their education: they should also be exposed to ideas which they would not think of for themselves. So I try to meet both of these goals by exposing them to a lot of new things and by allowing them to pursue their pre-formed interests."

Results. Although these students were Juniors and Seniors with two years of successful college experience behind them there was a widespread feeling of discomfort at the unaccustomed task of having to work in an unstructured situation, i.e., without the teacher's direct supervision; in the teacher's estimation this was true of about half the students. Some study groups "seemed to be exceedingly dependent at first; it took me a long time to cut the umbilicus there, but it seemed successful eventually. I wasn't aware that it had been successful until almost time for the report to be ready, so how sudden or gradual a process this was I don't know."

When the "guided" groups gave their reports, which came first, the teacher was deeply impressed with their quality. Having used a similar method before, but without any guidance at all, he had some basis for comparison. From his own account, he was outspoken to the class in his approval of the first reports. In his opinion, virtually all of the reports were of equally high quality. This result might be interpreted as meaning that the extra help received by the guided groups was unnecessary, since the unguided groups did as well. But the instructor felt that the oral reports of group projects were far superior to those given in previous semesters, in which no guidance was given. He attributed this to the combined effect of the model provided by the guided groups and of the generous approval he gave to the first reports. From the students' written comments at the end of the course, it seems that they shared the instructor's high opinion of the group reports, feeling that their own group work had been valuable as well as the reports given in class by the other groups. This was unusual, since in previous semesters the instructor had found that students looked with disfavor on reports by their colleagues, preferring to spend their time in class listening to the instructor. The favorable reactions of the students to the course are all the more remarkable in light of the fact that they felt the course required more work than was fair for a 2-credit course, an opinion shared by the instructor.

It may well be that a small amount of guidance given early, coupled with support and approval at a critical moment when students were quite unsure of themselves, were the key factors in the success of this course.

On the economic side, it should be added that the instructor estimated that the procedures followed in this course had saved him a certain amount of time. Although his conferences with the groups took some time, this was more than compensated by the fact that these group conferences reduced the number of individual contacts with students, as well as the time saved by reducing the number of lectures. In short, the teacher was able to see almost the whole class in small groups in less time than sporadic contacts with a smaller number would have taken in more conventional procedures.

General Psychology

This course, taught by the individual who taught the Culture and Personality course described above, was the second semester of General Psychology. At registration, the students were told that this was to be an experimental section, open only to students who had received a grade of A or B in the first semester of General Psychology. As it turned out, some students registered for this section who had lower grades.

Instead of three lectures per week, the customary number for this course, only one was given. In the instructor's opinion this was a better prepared, more abstract, and more highly compressed lecture than is ordinarily the case for him.

The students were organized into study groups of 3-5 students; with 66 in the class, this meant 15 groups. The only function of the groups was to study the subject matter of the course in whatever manner they saw fit. The teacher scheduled meetings with half of the groups and gave them considerable guidance as to how to go about studying together; other groups did not receive any more help than they asked for, which was almost nil, although the instructor made it clear that he was available and kept regular office hours including the regularly scheduled meeting time of the class.

On the final examination students who had been in the study groups receiving less guidance did slightly better than the others. The reaction of this class to self-directed study was not so uniformly favorable as in the Culture and Personality course. The instructor felt that this was largely due to the difference in the composition of the classes, since the General Psychology was largely Sophomores while the Culture and Personality class was largely Juniors and Seniors. Another possibility, of course, is that because the function of the study groups here did not require them to produce a finished product, such as the class report, many of them discontinued meeting; not only did the students lose whatever benefits may derive from working in small groups, they also had the unpleasant experience of being a member of a group that was falling apart.

Although there was no Control group with which a legitimate comparison could be made, the instructor, drawing on his past experience in teaching the same course, felt that the students had done quite well in the course, as reflected in the final examination.

He feels quite strongly that all students should eventually be exposed to self-directed study, since developing the ability to engage in it is one of the prime goals of higher education. But he also feels that the process of selection involved in the large number of students dropping out in the first two years of college means that self-directed study can be employed with a more able group in the last two years. For Freshmen and Sophomores he would not abandon self-directed study but would develop training methods to prepare them for it.

Independent Study in the Honors Program

The authors' original intention was to make arrangements for some students to do a large fraction of their academic work in the form of independent study and to compare the development of these individuals with matched individuals not given this opportunity. With the help of the Director of the Honors Program 10 matched pairs were selected from among Juniors in the Honors Program. Juniors were selected because it was felt that they would be able to cope with a large amount of independent study and that it would be interesting to observe their work in the Senior year. Every student approached was eager to participate. Unfortunately, it proved impossible to arrange independent study for all of them, so the matched pairs experiment was not carried out. Moreover, some members of the faculty took a dim view of a student's devoting so much time to independent study, so that many of the programs worked out for individual students only involved a small fraction of the student's work for the semester.

Eventually, interviews were obtained from nine students who had done some work in Independent Study. These students received credit toward a degree, but no letter grade, only pass or fail. In some cases the students met with the professor regularly, weekly or bi-weekly, and in other cases only occasionally, at the student's initiative. Topics covered included: Central European History, The Labor Movement, Psychological Research, Ways of Philosophical Thinking, Greek Tragedy, and Shakespeare.

According to the teachers interviewed, independent study was a success in that it was at least as good as Honors courses (small seminars) and in some cases quite superior. They felt that the students worked hard, thought hard and profited enormously. One student, a Junior, conceived and conducted an experiment that was published in a professional journal. Clearly, the teachers were satisfied and hoped to continue offering independent study.

Student reactions to independent study ranged from quiet satisfaction to wild enthusiasm. For some who had been overly concerned with getting good grades an entirely new vista opened up. The private, face-to-face meetings with the teacher required and readily evoked thorough and thoughtful preparation. Students spoke of the deep satisfactions stemming from discovering things for oneself, achieving breadth and illumination, for the first time getting a "real education." Some students found themselves so involved in independent study that they were devoting disproportionately large amounts of time to it. They realized that this would probably reduce their chances of getting high grades in other courses, but they were quite unconcerned about this. For them independent study seemed to be of greater, less transitory value than high grades. For some it became clear for the first time just what the professor saw in his discipline that attracted him so. They discovered the world of intellect, discovered that they were quite able to function in it, and that they enjoyed it. For some this was a marvellous experience, all found it rewarding and all planned to take additional independent study.

These student reactions are probably somewhat more favorable than those described by Bonthius _et al_ (1957). Perhaps the differences can be attributed to the selection of students for independent study; alternatively, it may be attributed to the noticeable lack of opportunity for independent intellectual work in this institution, a matter that will be taken up in detail in a later chapter.

Part III

Experiments Conducted in 1958-59

On the basis of the previous year's results it was decided to try to arrange as many experimental courses as possible, in each case constructing a 4-part final examination along the lines of the one used in the course in Physical Optics. The guiding hypothesis was derived from the results of the experiment in Optics: the lecture method promotes learning of facts and the ability to solve simple problems following a set pattern, while self-directed study promotes development of the ability to solve more complex problems and to learn independently. In the chapters that follow this idea will be referred to as the pattern hypothesis.

After obtaining permission from the appropriate Deans, letters were sent to the Chairmen of all departments in the College of Arts and Science, College of Engineering, and School of Business. The letter summarized previous results, gave some idea of future research plans, and asked the Chairmen to pass the letter along to faculty members who might be interested in participating in self-directed study experiments. After considerable discussion, experiments were arranged involving courses in Chemical Engineering, Economics, English Language, Marketing, Political Science, Physical Science (General Education Program), and Psychology. Instructors in several other departments were interested in participating (Civil Engineering, English Literature, Mathematics, and Sociology), but adequate control groups could not be arranged for their courses.

The instructors involved in the 8 courses agreed to allow one hour near the beginning of the semester for the research group to collect background information from the students, and one hour near the end of the semester for a course evaluation questionnaire. In addition, they agreed to give the same final examination to Self-directed Study and Control groups, this examination to be in 4 parts, lying along a hypothetical continuum of complexity, as described above. To aid the instructors in devising the final examinations, they were given a dittoed description of the general scheme to be followed. One of the research group joined in the discussion with the instructors when the final examinations were being planned. His role in these discussions was simply to clarify the general scheme, not to propose questions for the final, a matter better left to instructors. The dittoed "General Scheme for Final Examinations" is reproduced in Appendix A.

The beginning-of-term questionnaire covered 15 variables: two measures of ability (vocabulary scale and problem-solving test); five characteristics of the student or his campus life (sex, age, year in school, type of living quarters, sorority/fraternity membership); and eight aspects of his social and personal background (father's education, mother's education, father's position in community, religion, state residence, size of community, size of high school graduating class, and summer occupations). With very few exceptions, the Self-directed Study and Control groups were comparable on the 15 variables. The exceptions will be indicated as each experimental course is discussed.

Possibly the most important part of the preliminary questionnaire were the two measures of intellectual ability. As there was not adequate time for a standard intelligence test, a very brief procedure had to be developed. This consisted of the vocabulary scale of the Army General Classification Test and a short, objectively scoreable test of problem-solving ability, items 38 through 46 of Form G of the Test of Critical Thinking (Dressel and Mayhew, 1954). In no case were there appreciable differences in intellectual ability between experimental and control groups.

In the chapters that follow the experimental courses conducted in 1958-59 will first be described separately. These descriptions will be followed by an analysis, based mainly on the end-of-term questionnaires, which applies to the groups of courses as a whole. Finally, there will be reported a study of students' thought-processes in the classroom, using students from these experimental and control group courses.

Chapter 9

Freshman English

Fall 1958

Two teachers were involved in this study. The teacher of the Control section was the woman who taught the Experimental section in the previous year's study of the same course (see Chapter 5). The teacher of the two Experimental sections was of the same sex, and of comparable age and teaching experience. The two teachers shared similar views of education, believing that students can profit from somewhat greater individual responsibility and more mature treatment than they usually receive in Freshman English courses.

Experimental Procedure

There were two Experimental sections, listed in the University catalog one scheduled for classes at 8:00 A.M., the other at 9:00 A.M., Mondays, Wednesdays, and Fridays. The Control section was scheduled to meet on those days at 10:00 A.M. Students are assigned to these sections in a haphazard manner, and no one connected with the registration procedure was aware of the intended experiment; there is, as a result, no reason to think the students in these sections differ in any systematic way. When compared on the 15 background variables described above, the groups were, in fact, very similar, with no statistically significant differences appearing.

Control treatment. The Control section met three times a week, as scheduled. Texts for the course were a writer's handbook intended for Freshman English courses and an anthology of prose readings. Students were required to make seven speeches, write seven themes and several papers critical of students' speeches, and read assignments in the texts. In correcting the students' themes, the teacher indicated the section of the handbook explaining the error of grammar, punctuation, or usage. Students were required to correct their errors. The teacher estimated that 40 per cent of class time was devoted to discussion of the principles of composition and grammar, 50 per cent to listening to speeches and criticizing them, and 10 per cent to writing themes in class or taking examinations. Every student was required to have at least two individual conferences with the teacher.

Experimental treatment. The Experimental section met twice a week regularly, the third meeting being optional for students feeling that they needed additional help. The texts, assignments of readings, speeches, and themes were the same as in the Control group. For the first two themes, corrections were made in the manner described for the Control group. For the next three themes location of errors was indicated, but it was up to the student to identify the error himself. For the last two themes, the location was not indicated: the student was told the kind of errors he had made, and asked to find and correct them. The teacher estimated that 33 per cent of class time was devoted to discussion of principles of composition, the remainder to listening to speeches and writing. Questions of mechanics were discussed only in response to students' questions. There were no required individual conferences, but the teacher was available on the student's request.

Final examination. All three sections took the same final examination. Part I consisted of an objective test of mechanics; Part II was an objective test on analysis of writing (identifying thesis sentences, etc.); Part III was the final class theme; Part IV consisted of the final outline and speech. The grading of Parts III and IV was done by the two teachers together; each paper or speech was evaluated by both teachers independently, the grade given being their consensus. It was not possible to arrange for blind grading of these examinations, although it would have been desirable for the individuals assigning grades to have remained ignorant of the identities (Experimental or Control) of the students.

Results

In Table 1 the results are given for all students taking the final examination. It can be seen that there are no very striking differences among the groups. The SDS groups are slightly superior on Parts I, II, and III of the final, and the Control group is superior on Part IV, but none of these differences is statistically significant. These results do not confirm the pattern hypothesis proposed above (p. III-1).

Table 1

Final Examination Performance in Freshman English

		Final Examination							
		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental 1	21	8.2	3.7	7.2	2.1	9.7	2.9	8.4	3.9
Experimental 2	22	7.9	2.8	7.6	3.3	9.8	2.3	8.6	2.8
Control	25	8.5	3.3	7.8	2.5	9.9	5.1	7.8	1.9

Note: In this table, means are given as error scores; a lower score represents a superior performance.

Basing their decision on the slight superiority of the SDS group, the two teachers drew the conclusion that the experimental method had produced somewhat superior results and planned to go on with it in future semesters.

Discussion

Given the very small differences described above, the safest conclusion here is that there was no difference between the SDS and Control groups. It will be remembered that in the previous year's experiment with this course, the experimental group's performance was appreciably better than the Control. Can we account for the discrepancy between the results in the two years? It may be that the relevant variable is the attitude of the teachers toward self-directed study.

In the first year's experiment, the Experimental teacher favored self-directed study while the Control teacher believed strongly in a disciplinarian, teacher-directed approach. In the second year's experiment, both teachers favored self-directed study. It seems probable that the two teachers in the second experiment conducted their classrooms quite similarly--that is, in a rather informal manner--so that the main difference between the two groups was merely that one met twice a week while the other met three times a week. In the first year, even though not so many classes were cancelled for the Experimental group, there was a marked difference in the two teachers' approahces. This interpretation suggests that the teacher's philosophy of education expresses itself in many aspects of his behavior, not only in the decision to meet his class or cancel it, that it is the total philosophy underlying self-directed study that may make the difference, not merely the number of times the class meets.

In any event, the teachers were probably justified in drawing conclusions favorable to self-directed study. At the very least, it can be said that a Freshman English class could meet only two-thirds the usual number of times and be required to take a more than usually active part in teaching themselves to use English well, all without any great harm befalling them.

Chapter 10

Physical Science: Man in the Physical World

Fall 1958

This course, generally taken in the freshman or sophomore year, is one of several ways of satisfying the physical science requirement. It is a two-semester course with no prerequisites, and is administered by the Division of General Education.

Procedure

Two sections of the course were involved in the experiment. The students in these sections learned of the experiment at the first meeting of the class. The self-directed study (SDS) section was taught by a younger, less experienced person with the rank of instructor; the control section was taught by an older, more experienced person with the rank of assistant professor.

The SDS section met three hours per week for the first few weeks of the semester, two hours per week for the remainder. The control section met for six hours per week throughout the semester. Of the time spent in formal class meetings, the SDS teacher lectured almost 100 per cent of the time; the control teacher devoted about 60 per cent of the time to lecturing, 40 per cent to class discussion.

The SDS class was divided into 6 groups of 5 or 6 members each, the teacher making some attempt to form groups that were equivalent with respect to high school background in science and mathematics. Each of the groups was responsible for carrying out five projects, two of an historical nature and three experimental. The students were given the impression that these projects combined would contribute to their final grade in an amount equivalent to one or two hourly examinations. Since there were four such "hourlies" as well as a final examination, this meant that the students were under the impression that the projects counted for very little. Actually, according to the instructor, they counted for even less. The experimental work was performed without direction by the instructor other than setting the objective and making apparatus available. Handing in of group reports was the only homework required. According to the instructor, one group did excellent work, four did generally satisfactory work, and one did poorly.

The students in the SDS group were urged but not required to do the problems at the end of each chapter in the textbook. In addition, they were given a set of extra problems in mechanics with the suggestion that they do the problems in their groups.

The control group received four one-hour examinations in addition to the final examination. They were required to hand in problems from the textbook assigned as homework.

Discussion

On the whole, the results of this experiment are ambiguous. The Control group did slightly better than the SDS group, but the difference was probably not commensurate with the additional expenditure of the teacher's time. In addition, at the end of the semester the students in the SDS group showed a more favorable attitude toward independent intellectual work then did the students in the control group. This result will be discussed more fully in a later chapter. Perhaps the best commentary on this experiment is the SDS teacher's reaction that the experiment was a failure because the SDS group's performance was not *superior* to the Control group's, but that he planned to go on with an improved version of the SDS method of teaching.

Chapter 11

Principles of Economics

Fall 1958

This course, which is not open to Freshmen, is required for all students majoring in Economics; it is also one of the ways of satisfying the social sciences requirement for all students in the College of Arts and Sciences.

Experimental Procedure

In the semester of the experiment there were nine sections of this course. Two of them were involved in the experiment, the teachers of both of these sections being senior members of the Department of Economics, men with many years of teaching experience. During the semester, the research staff held several meetings with both professors and their teaching assistants. At these meetings the progress of the experiment was discussed and plans were made for the common final examination.

Control treatment. The Control group met three times a week for lectures. In addition, the students met in a small group tutorial once every two weeks. The student's grade in the course depended only on the final examination. The lecturer went to great pains to avoid covering any of the material in the textbook. Questions arising from the readings or from the lectures were taken up in the tutorials.

Experimental treatment. The SDS group was scheduled to meet three times a week. Instead, when the experiment was initiated at the beginning of the semester, the students were informed that the class would meet once a week. The student's grade in the course depended on the final examination, and--at the student's option--on two essay tests taken during the term to reduce the weight of the final for those who so desired.

There was no tutorial in the SDS group. Given the option, about two-thirds of the students elected to study on an individual basis, while the others formed study groups ranging from 4-10 students per group. According to the students' reports, of the 4 study groups initially formed, 2 never actually met to study, one group was quite effective, and one group began well but became progressively less effective and eventually decided to discontinue meeting. Attendance at the weekly lectures averaged about three-fourths of the enrollment; although office hours were set aside for individual conferences between the students and the instructor or his assistant, almost no use was made of the opportunity.

Of 80 students registering for the Control section 74 completed the course, plus 2 who transferred from the SDS section. Of the 77 students registering for the SDS section, 9 dropped the course, 4 transferred to a different section, and 2 withdrew from school. As will be discussed below, there is reason to believe that this difference in attrition rates between the two sections was, at least in part, a direct effect of the experiment.

Results

By about mid-semester it became clear that the SDS group was not doing very well. At a meeting of the staff involved in teaching the course with the research group, the teacher of the SDS group gave a very unfavorable description of the progress of the experiment: poor attendance at the weekly lecture, many complaints, and generally low morale. As indicated above, quite a few students dropped out of this section. At the teacher's instigation, some of them submitted explanatory letters. One student, who dropped out after a month, wrote that he had a job taking about 6 hours a day of his time; in addition, "one of my courses requires at least two hours of review each night. My other courses .... each require a half hour to one hour each day .... Economics, being an experimental, independent study class, is naturally set aside as there are no daily incentives in the form of quizzes or discussions which are present in my other courses." This preference for well-structured courses with clearcut, periodic assignments has been voiced very frequently by students with whom self-directed study has been discussed.

The two groups received the same 4-part final examination, devised by the instructors according to the plan described above (p. III-1). Part I, a 60-item multiple choice test, was taken by all 9 sections of Principles of Economics, thus providing a basis for comparison broader than the one Control group actually arranged for the experiment. The comparisons of SDS and Control groups are shown in Table 1.

It can be seen that the Control group was superior on all 4 parts of the final examination. The difference was statistically highly significant for Parts I, II, and III. When the performance on Part I of all 9 sections of the course is examined, the inferiority of the SDS group is clearcut: its performance was the poorest of the 9, the median for all 684 students being 39 correct out of a possible 60, while the median for the SDS section was 35.

Table 1

Final Examination Performance in

Principles of Economics

		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	62	106.3	18.3	49.2	20.6	101.0	31.8	149.8	20.0
Control	76	116.6	16.6	68.3	20.3	135.7	24.7	159.3	30.5
t		3.40		5.46		6.94		2.18	
p		$<.001$		$<.001$		$<.001$		$<.05$	

As it happened, the Control group had a somewhat higher percentage of Juniors and Seniors than the SDS group. An examination of the results subdivided by classes indicated the same trend as described above: the Control group's performance on the final examination was equally superior for each class.

Discussion

In the opinion of the teacher of the SDS section, the experimental method had failed; his students had not learned as much about economics as they ordinarily do in his class. Hence he returned to the teaching methods he had used in the past with this elementary course. However, his observations led him to believe that self-directed study might work well in elective courses, where the students were already well-grounded in the subject. Then, he felt, the student would be able to study on his own and undertake a significant project with every prospect of profiting from the experience.

In trying to understand the fact that the SDS group fared so poorly in this course, three possibilities come to mind. First, the introduction to the subject of economics at the university level may require more help from the teacher than many other subjects. Secondly, the students in the SDS were not only given fewer lectures, they also missed the bi-weekly small group tutorials in which the Control group participated. A third possibility has to do with the instructor's philosophy of education. In the course of these investigations, the authors formed a strong but largely undocumented impression that this may be a very important variable influencing the effectiveness of any teaching method. A teacher announcing that the student is to be more than usually responsible for his own education may create an impression that the student must "sink or swim," or he may emanate a kind of personal warmth that conveys the feeling that if help is needed it will be provided in a friendly, non-penalizing way. Not every offer of help conveys the latter impression: it is possible to offer help in a way suggesting that a student, by accepting the offer, will be judged as weak or helpless. McKeachie (1954), among others, has suggested that the teacher's behavior in this regard may be a very important aspect of the teaching process. It is easy to see how this variable might become critically important in a situation, such as the present study, where the student is informed that he is going to be largely responsible for teaching himself a wholly unfamiliar and perhaps a somewhat forbidding subject.

Chapter 12

Principles of Marketing

Fall 1958

This course, given in the School of Business and prerequisite to all other courses in marketing, is generally taken in the Junior year.

Experimental Procedure

Four sections were involved in the study, two experimental taught by one teacher and two control taught by another teacher. All groups used the same text; all had two tests during the semester, part essay and part objective. The SDS section met, on the average, two hours a week; the Control section met three hours every week. In class, both teachers avoided covering material that was in the textbook, instead emphasizing analytical and creative approaches to problems. The SDS teacher estimated that half of each class period was devoted to lecture and half to discussion; the Control teacher estimated 95 per cent lecture and 5 per cent discussion. Both teachers required their students to carry out projects concerned with the whole process of designing and marketing a new product. Thus the major differences between the two groups were in the number of class meetings per week, two for the SDS group and three for the Control, and in the amount of discussion occurring in class.

At the beginning of the semester it was agreed by the instructors that all sections would take the same final examination, divided into four parts according to the specifications described above (p. III-1). It should be mentioned that, at the outset, there was a fifth section, taught by still a third instructor; at the end of the term, he decided against giving his section the agreed-upon examination. This was unfortunate, since his group was quite different from the others: his lectures stressed factual coverage of the content of the course, and his group had no projects to do. As can be seen from the description of the teaching methods used by the groups that did take the same final examination, both the SDS and the Control group had experiences of a more self-directed nature than did the control group that de-materialized at the end of the term.

Results

The results of the final examination are shown in Table 1; the two SDS groups have been combined, as have the two Control groups. The Control group was significantly superior on Part III; none of the other differences approached significance.

The teacher of the SDS section felt that his students had not gained as much of the technical vocabulary of the field of marketing as they would have done with greater opportunity to use it in class. Apart from that, he was well enough satisfied with the results to be willing to continue with self-directed

Table 1

Final Examination Performance in Principles of Marketing

		Final Examination							
		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	65	46.8	5.4	27.8	3.1	33.7	4.2	18.9	4.1
Control	65	48.2	5.8	27.3	3.9	38.2	6.4	18.2	4.9
t		1.42		0.81		4.75		0.87	
p		n.s.		n.s.		$<.001$		n.s.	

study. Modifications he contemplated included increases in reading assignments beyond the textbook, and increases in the number of papers required, both with the intent of increasing the student's responsibility for educating himself.

Discussion

There is no clear evidence for the superiority of either group here. Although the Control group was distinctly better on one part of the final, the SDS group was slightly better on two parts. On balance, the Control group's performance was better, but the difference was quite small. The reduction from three to two contact hours per week did not prevent the experimental group from assimilating the main contents of the course, and it may well be, as the SDS teacher suspected, that the opportunity to study the subject independently in itself compensated for any small losses of subject matter. For example, at the end of the semester the experimental group showed a more favorable attitude toward independent intellectual work. This result will be discussed more fully in a later chapter.

On the other hand, the fact remains that the Control group's performance was slightly better, and moreover, that the pattern of results offers no support to the pattern hypothesis (see p. III-1).

Chapter 13

Political Science: American Government

Fall 1958

This study dealt with the first semester of a two-semester course in American Government. Although designed for Freshmen, it is open to all students and is one of the ways of satisfying the requirement that all students take a year of social sciences.

Experimental Procedure

Two teachers of comparable rank and experience participated. One taught a self-directed study section and a control group (Control Group A); the other taught a second control group (Control Group B). Students had no knowledge of the experiment when registering for the course. All three sections were assigned the same readings and took three one-hour essay tests during the semester. As scheduled, the two control groups met three hours a week, while the SDS group met only one hour a week. For both Control groups the teachers estimated that they lectured about 75 per cent of each class period, allowing the rest of the time for discussion; the teacher of the SDS section estimated that he allowed that group somewhat more time for discussion, 60 per cent lecture and 40 per cent discussion.

The SDS group was required to do two projects: field work with a political party and attendance at Boulder City Council meetings. Each students was required to submit two papers, one about each project. Control Group A had no projects; Control Group B was required to do one book report and one field trip. All three sections took the same four-part final examination, as described above (p. III-1).

When compared on 15 background variables, the three groups proved to be very similar, except that Control Group B had significantly fewer members of sororities and fraternities.

Results

The results of the final examination are shown in Table 1. The Control groups did better on Parts I, II, and III of the final; the SDS group did better on Part IV (learning new material). Comparing the SDS group with Control group A, taught by the same instructor, only the difference on Part II in favor of the Control group was statistically significant ($t = 2.00$, $p < .05$). These results give some support to the pattern hypothesis (see p. III-1).

Discussion

The teacher of the experimental section was not entirely satisfied with the results. He felt that the term papers were superficial in style and thought. He thought that the students might have done better work and gotten more out of the course if they had been given some guidance in what to look for in observing

Table 1

Final Examination Performance in Political Science

		Final Examination							
		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	49	15.7	5.5	12.7	5.9	15.6	5.2	11.3	4.9
Control A	58	16.4	5.4	16.7	10.1	16.4	3.1	10.3	5.3
Control B	58	18.4	5.2	15.0	4.7	17.5	4.0	10.3	6.4

political processes. He objected to the rigid schedule of one lecture a week, suggesting that it might be better if the instructor had simply used his judgment to cancel classes when, in his estimation, the students were prepared to do a block of work independently. In the following semesters, he modified his own teaching methods accordingly, preserving some of the features of the experimental method.

It should be mentioned that the original plan for this experiment involved another self-directed study section, taught by a third instructor. Just before the semester began, however, this individual decided that he did not wish to reduce the number of class meetings. Instead, he met his class three times a week, but he conducted his class quite differently from the Control groups described above. He estimated that only 30 per cent of class time was devoted to his lecturing, the remainder--70 per cent--going for discussion. These discussions generally involved a kind of Socratic technique in which one student at a time would be singled out for a series of questions. This group, which was intellectually equivalent to the other groups, gave the best performance on all four parts of the final examination. Whether this result was due to a difference among the instructors or to the methods employed cannot be elucidated. As a matter of fact, these two alternatives are not necessarily distinct, since it probably takes a rather demanding person to conduct a Socratic inquiry in a large classroom.

To sum up, these results give some support to the pattern hypothesis: although the self-directed study group was inferior to the Controls on Parts I, II, and III of the final, it was slightly superior on Part IV.

Chapter 14

Principles of Chemical Engineering

Fall 1958

This course, taken by Juniors and Seniors, is concerned with the study of fluid mechanics and heat transfer.

Experimental Procedure

The same teacher taught two sections, one experimental (SDS) and one control. The Control group met for 4 hours per week while the SDS group met for only 1 hour per week. The teacher estimated that in the Control group 50 per cent of each class meeting was devoted to his lecturing and 50 per cent to discussion; in the SDS group, he estimated, 90 per cent of the class meeting was devoted to his lecturing and 10 per cent to discussion. All in all, then, the Control group heard somewhat more than twice as much lecturing as the SDS group, in addition to which the Control group participated in considerable class discussion under the teacher's tutelage.

Both groups were assigned the same reading from a textbook in chemical engineering, and the same large number of problems. In the selection of problems and in his general conduct of the course, the teacher puts considerable emphasis on learning to solve new problems rather than on application of standardized methods.

Both groups took the same five two-hour tests during the semester. For these tests, the SDS group met with the Control group so that the tests would not conflict with the lectures given to the SDS group. Both groups took the same final examination. Because the instructor deemed that the entire 4-part scheme (see Appendix A) was not appropriate to this course, the final examination was designed to measure achievement at only two levels of complexity: first, simple applications of facts and principles, and secondly, more complex applications of facts and principles. To facilitate comparing the results of this study with other studies conducted in the academic year 1958-59, the two parts of the final will be referred to as Part II and Part III, respectively (i.e., Part I and Part IV are missing from the 4-part scheme).

Although the teacher graded the final examinations himself, he attempted to eliminate any possible bias by shuffling the papers of the two sections into one randomized pile. While grading, he avoided looking at the names.

There were 10 students in the SDS section and 17 in the Control section. The two groups were quite comparable on 15 background variables, none of the differences being statistically significant at the 10 per cent level. The SDS group was slightly inferior on the tests of intellectual ability. Two members of each group failed to take the final examination.

Results

The SDS section did at least as well as the Control on all five two-hour tests. The results for the final examination are shown in Table 1. It can be seen that the SDS group was slightly superior on Part II (simple) and markedly superior on Part III (complex). The fact that the greater superiority occurs on the more complex material offers some support for the pattern hypothesis (see p. III-1).

Table 1

Final Examination

		Part II		Part III	
	N	M	SD	M	SD
Experimental	8	31.9	6.1	45.6	8.1
Control	15	28.0	12.8	36.7	8.9
t		0.95		2.29	
p		n.s.		$<.05$	

Discussion

It was the teacher's impression that the self-directed group was very successful. All but one of the students in the experimental section worked hard, one or two of them becoming quite stimulated and probably doing much better than they would have under usual class conditions. In the teacher's opinion, certainly two and perhaps as many as seven of the students in the Control group would have benefited a great deal from participation in self-directed study. In spite of considerable floundering on the part of the SDS group early in the semester, they had been able to overtake and surpass the Control group. He planned to adopt some of the features of self-directed study in future courses.

The reader will have noticed that the Control group had 17 students and the SDS group only 10. There is some evidence that small discussion groups are more effective than large ones, but for the lecture method there is no clear difference (McKeachie, 1960). If the teacher had encouraged discussion in the SDS group, the suggestion would be plausible that size of class rather than self-directed study was the effective variable. As it happened, the SDS class time was almost all devoted to lecturing, while the somewhat larger (but not very large) Control group had considerable opportunity for discussion as well as lectures. On balance, although the question cannot be settled, it seems more likely that self-directed study was the effective variable.

It is interesting to remark that the instructor might have divided the entire group of 27 students into 3 small groups each meeting 1 hour per week, exactly like the SDS group: the total time he would have spent in class, 3 hours per week, would still have been 1 hour less than the time he gave to the Control group alone. Thus, if smaller classes are called for, they could be arranged without extra expenditure of faculty time by reducing the number of hours each student spends in class, i.e., by placing greater emphasis on self-directed study.

Chapter 15

General Psychology - I

Fall 1958

This is a two-semester introductory course, one of the ways of fulfilling the requirement that all students in the College of Arts and Sciences take a year of biological sciences. The course is most often taken in the sophomore year, but any student may enroll in it except Freshmen scoring below the 70th percentile on the scholastic aptitude test given to all entering Freshmen. There are usually about 1,000 students in the course.

Experimental Procedure

Of the fourteen sections offered in the Fall of 1958, three were designated as experimental (SDS) and three were designated as control. The SDS sections had different instructors and slightly different procedures: they met from about 9 to 16 hours during the semester (i.e., weekly or biweekly), had one or two tests (essay or combined essay and multiple choice), and they were required to do a term project. The Control sections met for 48 hours (i.e., three times per week), had ten multiple-choice tests, and were not required to do a term project. According to the teachers' estimates, class time in the SDS groups was divided about evenly between lecture and discussion, while in the Control groups about two-thirds of the time was devoted to lecturing, one-third to discussion.

All groups took the same 4-part final examination, constructed according to the scheme described above (see p. III-1), designed to measure performance at four different levels of cognitive complexity. All items were objectively scoreable. At the beginning and end of the semester students in these sections took the test battery described above (pp. III-1 ff.), providing measures of intellectual ability, background information, and end-of-term course evaluations as well as a measure of value placed on academic independence. Presentation of the results for the latter two variables will be deferred until Chapter 18, where results for all experimental courses in 1958-59 will be considered together.

Posttest. In May of 1959, fifteen months after the course was over, a follow-up study was conducted. To ensure that the test would be administered to students who had learned an appreciable amount in the course, only students at or above the median in total final examination performance were considered. Of these, a random sample of 120 were invited to participate. As it turned out, only 84 of those invited were still at the University, and 74 of these took part in the delayed posttest. These 74 were similar to the population from which they were drawn (students above median on final examination) on the 15 background variables; moreover, those who had been in the control sections were similar to those who had been in the SDS sections.

The letter of invitation gave no inkling that the research they were asked to help us with had any connection with a particular course, and when the students took the tests they were urged not to reveal anything about them to their friends. All respondents were asked if they had suspected in advance the purpose of the research, and none of them indicated that they had. There is, consequently, no reason to believe that any of them were prepared for the surprise that greeted them.

The delayed posttest was designed to test psychological knowledge and acumen at both ends of the continuum of cognitive complexity represented by the 4-part final examination.

The first part of this delayed posttest consisted of 20 multiple-choice items drawn from Part I (factual information) of the final examination, the questions used being exactly the same as on the final. In selecting questions for such a delayed posttest it is important to use items of an appropriate level of difficulty. Items that were universally forgotten, for example, would have no discriminating power, since good and poor students alike would fail them. Accordingly, on the basis of an item analysis of the final, the posttest was constructed so as to contain many items that were quite easy on the final, on the assumption that fifteen months of forgetting would bring those items to a useful level of difficulty. This procedure, which was quite successful, has been described in more detail elsewhere (Gruber and Weitman, 1961).

In the second part of the delayed posttest the students were presented with a report of a fictitious, faulty experiment and asked to comment on it. These comments were later rated for completeness and adequacy of criticism, overall grasp of the material, and inclusion of constructive suggestions. Both the rating-rerating and inter-rater reliability of the rating scheme were satisfactory (N = 20, Pearson $\underline{r}$'s = .94 and .87, respectively).

After completing the second part of the posttest the students turned in their papers and were given the third part. This was a set of 10 true-false questions about the fictitious experiment they had just read, designed to test the students' grasp of the major weaknesses of the research design and understanding of the errors in the conclusions that the fictitious experimenters had drawn. Thus, the second and third parts of the posttest involved learning and application of new material with special emphasis on critical thinking.

Results: Final Examination

Analysis of variance. With the large number of students involved in this course, a meaningful study of the interactions among certain independent variables was possible. Of particular interest is the possibility of an interaction between treatment and ability, linked as it is to the hypothesis that the abler student may profit from self-directed study but not the less able ones. Similarly, an interaction between treatment and year in school would bear on the hypothesis that students are not ready for self-directed study until they have been in college a year or more. Sex was used as one of the independent variables because it has so often been found that male and female students do not react alike to educational treatments. For example, it was quite possible in the present study that men and women would react differently to the mere fact of innovation.

Two of the independent variables were automatically arranged in two levels each: treatment (SDS, Control) and sex (male, female). Since there were very few students in the course except Freshmen and Sophomores, year in school was arranged in two levels, Freshman and Sophomore plus. Almost all of the latter were actually Sophomores. Ability scores were a composite based on the vocabulary scale and the Dressel test (see p. III-1). These scores were divided into five quintiles containing as nearly as possible equal numbers of cases.

For each dependent variable, such as performance on Part I of the final examination, three three-way analyses of variance were carried out, one for each of the following combinations of independent variables: treatment x ability x year in school (TAY), treatment x sex x year in school (TSY), and treatment x sex x ability (TSA). This procedure was followed rather than one four-way analysis chiefly because the latter would have resulted in a certain number of empty cells. As it happened, the three-way analyses were easier to program in FORTRAN for use with the computer then available to us.

Results for four dependent variables, the four parts of the final examination will be presented in this chapter. Four other dependent variables dealing with the students' evaluation of the course and their attitude toward independent intellectual work, were handled in the same manner (three three-way analyses of variance per variable), but results for these will be deferred until Chapter 18, dealing with those variables for all studies conducted in 1958-59.

It will be noted that the three analyses performed for each dependent variable are overlapping. Treatment occurs in all analyses, and ability, year in school, and sex occur in two analyses each. The important interactions, treatment x ability and treatment x year in school, occur in two analyses each. In most cases, the results from these repeated analyses of the same effects yielded similar significance levels. When there was a discrepancy involving a dependent variable having one degree of freedom (treatment, year in school, and sex), a *t*-test was used. For comparisons involving ability, having four degrees of freedom, the Duncan Test (Duncan, 1951) was always used in order to arrive at a compact decision as to the way in which ability levels differed in performance. In using the Duncan Test it is necessary to use the Within Mean Square (error term) of the analysis of variance. In all cases involving discrepant *F*-ratios, the WMS from the analysis yielding the nonsignificant *F* was utilized. It is felt that these procedures represent a reasonably conservative approach to the handling of multiple comparisons.

Final examination performance. The statistically significant results for the analyses involving the four parts of the final examination are summarized in Table 1. In this table only significant effects have been entered. The *F*-ratios associated with these effects are shown in Table 2.

The mean performance on the four parts of the final is shown in Table 3. It can be seen that the mean for the Control group is higher for all four parts; the differences on Parts I and II are highly significant, but the differences on

Table 1

Summary of Analyses of Variance of Performance in General Psychology

(Only significant effects included)
N = 292

Source of Variation	Final Examination Part I	Part II	Part III	Part IV
Treatment	Control higher	Control higher	-----	-----
Ability	Level 5 higher than rest	Level 5 higher than rest	Levels 3,4,5 higher than Levels 1,2	Levels 5 and 2 higher than rest
Sex	-----	Males higher	-----	Females higher
Year in School	-----	-----	-----	-----
Treatment x Ability	-----	-----	-----	Complex interaction
Treatment x Year	-----	-----	-----	Fresh. Control higher than Fresh SDS. Sophomores intermediate

Note: The students were divided into 5 ability levels, level 5 being the highest 20 per cent, etc.

Table 2

F-ratios Obtained in Analyses of Variance

N = 292

Source of Variation	df	Analysis	Final Examination Part I	Part II	Part III	Part IV
Treatment	1	TAY	21.96***	19.74***	0.00	1.67
		TSY	16.97***	11.56***	0.00	1.67
		TSA	23.08***	10.95***	0.00	.00
Ability	4	TAY	7.52***	7.22***	2.73*	5.00*
		TSY				
		TSA	8.27***	8.10***	2.50*	4.12**

(Table 2 continued on next page)

Table 2 (continued)

Source of variation	df	Analysis	Final Examination Part I	Part II	Part III	Part IV
Sex	1	TAY				
		TSY	2.09	2.50	.71	5.00*
		TSA	6.94**	5.24*	.50	3.53
Year in School	1	TAY	.06	.42	2.27	0.00
		TSY	.17	.31	2.14	0.00
		TSA				
Treatment x Ability	4	TAY	.30	.21	.45	2.78*
		TSY				
		TSA	.60	1.19	.50	1.18
Treatment x Year	1	TAY	.47	2.12	1.36	6.11*
		TSY	1.04	.94	.00	4.17*
		TSA				

Note: TAY refers to analysis involving treatment, ability, year in school.
TSY refers to analysis involving treatment, sex, year in school.
TSA refers to analysis involving treatment, sex, ability.

* Significant at .05 level of confidence
** Significant at .01 level of confidence
*** Significant at .001 level of confidence

Parts III and IV are negligible. This result, while it indicates a clearcut superiority of the Control group at the time of the final examination, also provides some support for the pattern hypothesis (see p. III-1), since the difference is marked for tests involving the simpler cognitive tasks and virtually nil for tests involving the more complex cognitive tasks.

One of the instructors involved in this study taught a Control and an SDS section, both of which had about 40 students, as compared with the other sections which had 125-150 students. In this experiment-within-an-experiment, involving the same teacher for both groups, the pattern hypothesis was somewhat more strongly supported, since the SDS group was actually superior on Part IV of the final, and the difference in pattern of performance between Control and SDS was highly significant. In addition, this teacher, who had been rather skeptical

Table 3

Final Examination Performance in General Psychology

N = 292

		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	155	24.5	4.86	8.8	2.51	3.6	1.58	4.0	1.50
Control	137	27.3	4.13	10.0	2.24	3.7	1.62	4.1	1.45

about the value of self-directed study, was delighted with the quality of the term projects done by the students in the SDS group. All in all, then, he felt that self-directed study was definitely the superior method, in that the SDS students had independently learned about as much as the Controls had learned under a strict regimen of lectures and tests, in addition to which the SDS group had done worthwhile research projects.

Intellectual ability and self-directed study. In general, as indicated in Table 1, the more intelligent students did better on the final examination. With one exception, however, there was not a significant interaction between ability and treatment. In other words, the results do not confirm the hypothesis that self-directed study might be particularly effective for the bright student and particularly dangerous for the dull student. The one exception referred to above involves a complex interaction between ability and treatment with respect to performance on Part IV of the final. To examine this result in further detail, the students were subdivided in quintiles according to ability and further subdivided by treatment, yielding 10 sub-groups which were then ranked in order of mean performance on Part IV of the final. The 5th quintile (most able) SDS students ranked highest on Part IV, while the 5th quintile Controls ranked 5th on Part IV; this suggests that the most able students did profit in some measure from self-directed study. But for the 4th quintile the relation is reversed, 4th quintile Controls performing significantly better than 4th quintile SDS subjects (ranked 2nd vs ranked 8th). Among the least able students (1st quintile) the SDS group was ranked 9th and the Controls ranked 10th. Thus, although there was a statistically significant interaction between treatment and ability on Part IV, it is far too complicated in detail to allow any obvious interpretation. The best that can be said is that it gives some support to the idea that very bright students gain from exposure to self-directed study, but no support to the idea that the least intelligent students suffer from it.

Treatment and year in school. Here too the results are essentially negative, except for Part IV of the final. On Part IV there was a significant interaction between treatment and year in school. Freshmen enrolled in the SDS group were significantly inferior to Freshmen enrolled in the Control group. The performance of the latter was slightly higher than performance of Sophomores. Sophomores' performance on the final examination was not affected by the treatment. In sum, the results give some support to the hypothesis that self-directed study will be more disruptive for Freshmen than for other students. This statement should be qualified in two ways: first, this was only true for Part IV of the final; secondly, as will be discussed in Chapter 18, Freshmen in self-directed study may have gained the most in growth toward a favorable attitude toward independent intellectual work.

Results: Delayed Posttest

Table 4 shows the results for the delayed posttest, given 15 months after the course ended. Included with these results are the final examination scores for these students on the same 20 items that were used as the test of retention of factual knowledge.

Table 4

Final Examination and Posttest Performance of Students in General Psychology

	Experimental N = 41		Control N = 30	
	M	SD	M	SD
Factual knowledge				
Final examination	14.7	1.76	16.1	1.88
Posttest	12.9	2.46	12.4	2.50
Learning and evaluating new material				
Open-ended	6.61	3.04	6.47	2.43
True-false	7.90	1.46	6.97	1.35

It can be seen that the Control subjects were somewhat better on these items at the time of the final examination, but at the time of the posttest the positions are reversed. The SDS group is very slightly superior on the open-ended part of the posttest, dealing with learning and evaluating new material, and they are noticeably superior in their responses to the true-false questions dealing with the new material. The two scores based on responses to new material were summed and the resulting frequency distributions inspected. The chief difference between Control and SDS subjects was the higher proportion of superior performances in the SDS group. The difference between distributions just failed being statistically significant when evaluated by the Mann-Whitney U test ($U = 746$, $z = 1.54$, $p < .06$).

Table 5 shows in more detail the differences in retention between the two groups; it is clear that the Control students forgot more.

One peculiar result was accidentally obtained. The delayed posttest had a space for the student to fill in the name of the instructor he had had in the course (although, of course, that information was already on record). Of the

Table 5

Retention Fifteen Months After Termination of General Psychology

	Experimental	Control
	N = 41	N = 30
Gain or no loss	10	2
Small loss (1-2 points)	16	7
Moderate loss (3-4 points)	11	11
Large loss (5-9 points)	4	10

Chi-square = 10.03, df = 3, $p < .02$)

Control students 27 per cent had forgotten their instructor's name, while in the SDS group only 9 per cent had forgotten. The difference is statistically significant (chi-square = 3.97, df = 1, $p < .05$). This result is all the more surprising when it is considered that the Control students saw their instructor from three to five times more often than the SDS students saw theirs.

Discussion

These results illustrate how important it is to consider not only how much the student has learned, but also *how* he has gone about learning it. The Control group was clearly superior at the time of the final examination, but in the course of 15 months they forgot more, so that on the delayed posttest the SDS group was somewhat superior. This effect may perhaps be likened to Kersh's laboratory study (1958) mentioned above (p. 2-1). In this experiment subjects receiving careful guidance from the experimenter performed better just at the end of the training period; subjects in the "independent discovery" group were still baffled by the problem, but they had become engaged with it and spontaneously continued to think about it, with the result that their retest performance two weeks later was superior to the well-guided group.

The confused groping that characterizes many students in the self-directed study situation, especially at first, does not compete well with the smooth, well-practiced behavior of students who know how to "case" a teacher by listening to him for 40-50 hours. Instead, self-directed study initiates two struggles on the student's part: first, he struggles with himself to develop his ability to study independently; then, if he has been successful, he can struggle with the subject matter itself. This process, clumsy as it sometimes seems, may have effects more permanent than those produced by listening to a lecturer who is far more diligent than his students.

Chapter 16

General Psychology - II

Fall 1958

This course is the second semester of the one year sequence in General Psychology. Most students take the first part of the course in the Fall and the second part in the Spring. But there is considerable demand for so-called "trailer" courses, with the first part in the Spring and second part in the Fall. Often, students registering for these out-of-step sequences have failed the course previously, dropped out of school for a semester, or transferred from another school. Consequently, students in the present study, which was such a trailer course are not altogether typical of students taking the course in the normal sequence.

Procedure

Two teachers participated. The teacher of the Experimental group (SDS) was a male and an assistant professor; the teacher of the Control group was a female and a graduate student. Although she was a graduate student, she had had several years of teaching experience, including the planning of a new course. Both teachers were probably very good lecturers, so that the disparity in rank does not necessarily reflect any great difference in competence to teach the course in question.

When students registered for this course they were given no indication that an experiment was planned. The two groups were fairly similar on the 15 background variables (see p. III-1). As compared with the Control group, the SDS group had more students from states other than Colorado, more males, and fewer dormitory residents, but none of these differences was significant at the 5 per cent level. On the remaining background variables, differences were negligible.

The SDS section, scheduled to meet three hours per week for lectures, had only one lecture per week; the other two hours were used for discussions in which participation was optional. According to the instructor's estimate, the lecture period actually involved a considerable amount of interchange between himself and the students, only about 50 per cent of the period being used for actual lecturing. The Control section met three times per week, as scheduled. According to the instructor, about 90 per cent of the period was devoted to her lecturing, the remainder for questions and discussion. The Control group students were required to do two small projects and/or papers (such as constructing a test or writing a paper on personality theory). The SDS group had no such assignments. Both groups had the same textbook and the same book of supplementary readings. The SDS group had a list of additional readings as well.

During the semester the SDS group was given three one-hour objective tests; the Control group was given three one-hour tests involving both objective items and short answer questions. A substantial part of the final examination taken by the two groups was the same, this part being divided into the four parts described above (see p. III-1). In addition, each teacher gave his own group final examination questions that were not taken by the other group; these will not be considered here. There is some reason to believe that the final examination may have been weighted with questions covering topics that the Control teacher had stressed in class. As a matter of principle, the SDS teacher felt that students involved in self-directed study should be able to do well in an examination constructed for a conventional course. As a result, the examination, rather than being the joint product of the two teachers, was largely constructed by the Control teacher and accepted by the SDS teacher.

Results

The main results are shown in Table 1. The Control group was significantly superior on all four parts of the final examination.

Table 1

Final Examination Performance in General Psychology - II

		Final Examination							
		Part I		Part II		Part III		Part IV	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	129	4.6	1.3	15.2	4.8	.9	.6	14.8	5.1
Control	137	5.1	1.6	19.2	4.3	1.7	.6	16.7	5.2
t		2.79		7.11		[a]		3.00	
p		$<.01$		$<.001$				$<.01$	

[a] For Part III the possible scores were the integers from 0-3, and it was not deemed advisable to use the *t*-test. In a test of the difference between the two distributions of scores, Chi-square = 36.5, df = 3, $p < .001$.

Several other results are worth mentioning here, although they will be taken up in Chapter 18 as well. Control students liked the course better than did SDS students. For women, the difference in one of our two course evaluation measures was statistically significant ($t = 3.35$, $p < .01$); for men, the corresponding difference was in the same direction but much smaller and not significant.

On the other hand, at the end of the semester women in the SDS group displayed a more favorable attitude toward independent intellectual work than did women in the Control group ($t = 2.59$, $p < .01$); again, for men the corresponding difference was in the same direction but much smaller.

Discussion

This study happens to be the only one in the whole program in which one teacher was male and the other female. As it happened, the Control teacher was a highly attractive young woman. It is of some interest to note that the correlation of final examination performance with vocabulary scores is noticeably higher for students of the same sex as the teacher than it is for students of the opposite sex. The mean correlations of all four parts of the final are as follows: SDS - male, r = .36; SDS - female, r = .11; Control - male, r = .10; Control - female, r = .35. On the whole, vocabulary and course performance are significantly correlated if the teacher is of the same sex, but not if the teacher is of the opposite sex. This result should not be treated too seriously: it was entirely unpredicted, it cannot be crossvalidated in the present study, and it is not borne out in the results of the Dressel test of critical thinking. Nevertheless, it is tempting to speculate that when sex enters the teaching situation it becomes a more important variable than intelligence.

In any case these differences in patterns of correlation do not affect the main result of this study (superiority of Control group): the superiority of the Control over the SDS group was true of both sexes.

The performance of the experimental group was obviously inferior to that of the Control group. Candidly, however, one is hard put to it to make even a reasonable choice of the critical variable. First, it may have been the difference in the sexes of the instructors. As indicated above, the sex of the instructor may well be relevant to the student's performance in the course; moreover, it should be noted that there are far fewer women teachers than men, so that a course with a woman teacher at the University is ipso facto more unusual than one with a man. Secondly, the critical variable may have been the fact that the Control teacher played a greater role in constructing the final examination. Thirdly, it may have been the fact that the Control students were required to do projects but the Experimental students were not, a reversal of the usual procedure. Fourthly, it may have been the unusual sample of students involved in this "trailer" course. Finally, of course, it may have been the reduction in contact between the instructor and students in the Experimental group.

Under some circumstances, it might be considered legitimate to ignore such a badly controlled study. But within the context of the present program this is not possible. The initial decision was to conduct as many experiments as possible, the sole criterion for inclusion of a course in the program being the reduction of contact hours in one group. In pursuing this plan, the authors were explicitly aware of the likelihood that experimental rigor would diminish as the breadth of the program increased; the plan was pursued precisely because it was felt that the gain in breadth would compensate for the loss in control. With these considerations in mind, therefore, the reasonable conclusion to draw from the particular experiment described in this chapter is that under some circumstances reduction of time spent in formal classroom teaching may lead to an impairment of the student's performance in the course.

The SDS teacher now believes that it is important to give more guidance to students at this level than he did in this course. Because he was involved in several courses in the Self-Directed Study Program (see Chapter 8) involving different degrees of teacher-direction, he has been in a good position to form a personal judgment about problems involved in self-directed study. He now feels that it is important to provide the students with some guidance on how to proceed with self-directed study. One valuable plan is to divide the class into small groups whose function it is to study together and to carry out group projects; this works especially well when the teacher takes specific steps early in the semester to help the groups plan their work and get started on it. As pointed out in Chapter 8, with classes of as many as 100 students it is possible for the teacher to meet once or twice with each separate group. With much larger classes, this procedure might involve a prohibitive amount of the teacher's time, but other, more efficient guidance procedures might be invented (see, for example, Chapter 2).

Chapter 17

Other Studies Conducted in 1958-59:

Mathematics, Chemical Engineering

This chapter presents briefly three studies of self-directed study courses in which no parallel sections were available. In each case, only the teacher's description of the course will be reported.

Mathematics: Analytical Geometry and Calculus - III

This course is the last part of a three semester sequence in Analytical Geometry and Calculus, each part receiving 5 semester hours credit, or nearly one-third the academic load carried by a typical student. The entire sequence is generally taken by all students planning to major in mathematics. The third part of the sequence deals with infinite series, partial differentiation, and multiple integration.

Although there were several sections of the course being taught in the Fall of 1958, it proved impossible to arrange an experiment involving a parallel section with comparable students. Instead, one teacher agreed to participate in the self-directed study program by teaching an experimental course which would be open only to students who had gotten a grade of A or B in the preceding course.

Ordinarily, this class would meet for one hour per day, five days per week. In the self-directed study group, the number of classes was reduced from five to three per week. The teacher estimated 50 per cent of class time was devoted to lecturing, and 50 per cent to problems. Four one-hour examinations were given during the semester, as well as the final examination.

Results. According to the teacher, the reduction in formal class meetings had mixed effects on the achievement of the group. The reduced practice and review with any given type of problem resulted in these students having somewhat less than usual proficiency in certain mathematical manipulations and operations. On the other hand, these students were better able to get to the heart of a problem. The teacher attributed this to the fact that they had been obliged to find out more for themselves than is usually the case. The teacher was well satisfied with the results, but he had serious doubts about using self-directed study with a less select group of students.

Chemical Engineering: Advanced Chemical Engineering Thermodynamics

This course, ordinarily taken in the second year of graduate work, presents a quantitative development of the fundamental laws of thermodynamics and their application to chemical processes. Considerable emphasis is placed on engineering applications. Instead of the scheduled three hours per week, this class met for only one hour. The teacher estimated that 80 per cent of the class time was devoted to his lecturing, 20 per cent to discussion. Two textbooks were used, one on thermodynamics, one on chemical thermodynamics. Problems were assigned, but homework was not collected. Two one-hour examinations plus the final examination were given, all of these consisting of problems. In addition, each student presented a talk in class on a specially assigned topic. There were 12 students enrolled in the course.

Results. According to the instructor, the students were initially dismayed at the lack of detailed instruction and at the level of original analysis required by the test problems. It was his impression that most of the students worked faithfully in the course, but that all of them were baffled and confused by parts of the course. In the instructor's words, "More classroom contact might have helped them, but I do think they benefited from having to dig the material out themselves." He felt that they had eventually adapted to the method of teaching. Moreover, in the following semester, most of the same group were enrolled in a course with the same instructor on kinetics. Success in this course depends on the students' having a background in thermodynamics; according to the instructor they all seemed adequately prepared for it. He added, "... in fact it was the most responsive class I ever taught."

Chapter 18

Educational Values, Intelligence, and Self-directed Study

The purpose of this chapter is to draw together the diverse strands of evidence bearing on the relations among self-directed study, the ability of the student, and his attitude toward different aspects of academic work. Presented first will be some results bearing on a change of attitude as a consequence of participating in self-directed study; in this section, educational treatment is the independent variable and attitude is the dependent variable. Secondly, we will examine the way in which intellectual and attitudinal factors contribute to the student's reaction to self-directed study.

Attitude Toward Self-directed Study

Evaluation of experimental courses. The course evaluation questionnaire was modified from an instrument used for similar purposes at Antioch College. Two general types of item were included: relative and absolute evaluations of the course. In the four relative evaluation items, the student was asked to compre the course with other courses he was taking during the same semester on five-point scales, rating it for general interest, motivation to work for the course, amount of factual material learned, and mastery of essential ideas. The five absolute evaluation items again dealt with interest, amount of factual material learned, and mastery of essential ideas, plus two items asking point blank whether the student had liked the teaching methods used in the course and would choose to be in another course like it.

Attitude toward independent intellectual work. It was felt that the course evaluation questionnaire, which tapped the student's direct reaction to the experimental course, should be supplemented by a somewhat more indirect way of assessing the student's attitude toward independent intellectual work, since specific reactions to the course might be colored by the student's immediate fortunes in the course, and would not necessarily reflect his general attitude toward educational methods. Accordingly, a test of attitudes toward educational methods was constructed. The student was asked to allocate the time a hypothetical student would spend on various academic activities. Instructions for the task were as follows:

> An imaginary student at C.U. is going to take five 3-hour courses ... (and) plans to devote 9 hours a week to each course, including class time. This amounts to 135 hours per course (for a semester). Consider the most representative of the five courses and assume you are not bound by University regulations. How would you recommend the student spend his time in order to get the maximum benefit from the course?

A list of 11 activities was provided, and the student was asked to indicate the recommended amount of time to be spent on each. The list was intended to cover the range of things students might do in connection with courses, allowing enough detail to elicit attitudes toward independent intellectual work but keeping the list reasonably short. On an a priori basis (i.e., before examining any results), five items were selected as activities involving "academic independence": reading materials other than the textbook, out-of-class discussions with students, out-of-class discussions with the teacher, writing term papers, and working on term projects. Two items were selected as definitely "spoonfeeding": listening to lectures and reading class notes. Four items were considered to be somewhat ambiguous in this regard and will not be considered in the present results, although most of them might conceivably qualify as teacher-dependent: participating in class discussions, reading the text, taking objective tests, and taking essay tests.

The vocabulary test was administered during the first week of classes, the questionnaire on course evaluation and educational attitudes during the last week of classes, and the final examinations at the regularly scheduled time after classes had ended.

Results. On the course evaluation questionnaire, all items were constructed so that the mid-point (of the scale from 0 to 4) was 2; since there were 5 absolute evaluation items, a score of $5 \times 2 = 10$ would represent a neutral response to the course, lower numbers a negative response, and higher numbers a positive response; and since there were 4 relative evaluation items, the neutral score here is $4 \times 2 = 8$. On the academic independence and spoonfeeding items, a scaling procedure was used to correct for the students' tendency to use round numbers. In the results shown in Table 1, the mean scores and standard deviations are given in these arbitrary scale values; in parentheses next to each mean is given the number of hours equivalent to that score.

Table 1

End-of-Term Course Evaluations and Attitudes Toward Independent Intellectual Work in Seven Experimental Courses

		Absolute Evaluation		Relative Evaluation		Academic Independence		Spoon-feeding	
	N	M	SD	M	SD	M	SD	M	SD
Experimental	396	10.16	3.86	7.72	4.02	7.98(40)[1]	4.05	9.19(45)	3.12
Control	475	11.01	6.49	9.00	3.19	6.53(31)	4.40	9.82(50)	2.79
t			2.39		5.12		5.04		3.11
p			<.02		<.001		<.001		<.01

[1] See text for explanation.

The results shown in Table 1 are for the group of seven courses on which all information was available for the matrix of correlations to be taken up in the section following this one. Table 1 can be considered as reflecting a posttest only design, since experimental and control groups are compared on tests given at the end of the semester, with randomization of subjects between groups.

Table 1 presents an odd paradox. On both course evaluation measures, students in the self-directed study groups evince an unfavorable reaction to self-directed study: the control students give more favorable evaluations than do the experimental students. On the relative evaluation measure, the self-directed study mean is definitely on the unfavorable side of the neutral score of 8, i.e., these students liked the experimental courses less than they liked other courses they were taking.

In spite of this unfavorable reaction, the experimental group displays a general attitude toward educational method which is significantly more favorable to academic independence and less favorable to spoonfeeding than do the controls.

The unfavorable direct reaction to self-directed study courses is borne out by a wealth of observation. In almost every such course, a crisis of morale occurs. This phenomenon has been well described by an investigator in another institution, L. R. Beach at Whitworth College. In his study, which was in many respects similar to ours, "...a decline in enthusiasm showed through as the groups became somewhat dissatisfied with their progress, the difficulty of the text, and the lack of structure to the course experience. ...this downward trend continued until the course was about two-thirds completed. Then there began a steady climb in general reaction to the challenge of the relatively independent study and close interaction with the discussion group. At the end, while there was still some frustration and some threat regarding the approaching final exam, feelings ... were again high" (Beach, 1961). In our experiments conducted at the University, reaction at the beginning and end was less favorable than Beach's description, but the crisis of morale is similar in kind. The differences between his students' reactions and ours may be attributable to the different methods he used, or to the difference in atmosphere prevailing in a small liberal arts college as compared with the College of Arts and Sciences in a large university (see Chapter 20).

One observation paralleling the present contradiction between course evaluations and expressions of general attitudes has been described in Chapter 6. Although the students in the experimental group found self-directed study generally distasteful, more experimental than control students enrolled in a continuation of the course. An interpretation of these findings will be offered in the concluding chapter of this report.

Various efforts were made to test the hypothesis that students in those self-directed study courses which were most successful in terms of final examination performance would have had a more rewarding experience, which would be reflected in more favorable course evaluations. None of these efforts produced confirmation of the hypothesis. The best opportunity to test it arose in connection with a course that was not included in the results shown in Table 1, because, as mentioned above, Table 1 is restricted to courses on which complete data for the correlation matrix (see below) were available. In Principles of Chemical Engineering (Chapter 14) the experimental group was distinctly superior to the control group on the finals and on five tests during the term; nevertheless, the course evaluations were much more favorable in the control group than in the experimental group, the mean differences being larger than those shown in Table 1. Thus, "successful" experiments do not seem to produce more favorable student reactions than "unsuccessful" ones. When we turn to the correlation of individual performance with individual course evaluations, however, a somewhat different picture emerges.

Intercorrelations among performance, ability, and attitudes

This section examines the intercorrelations among the parts of the final examination, the vocabulary scale, course evaluation, and attitude toward educational methods. To save space, correlations involving the Dressel test of critical thinking have been dropped from the table to be presented; in general, these correlations were lower than those involving the vocabulary scale but in the same direction. Likewise, correlations involving relative course evaluations are not included; in general, these correlations were lower than those involving absolute evaluations but in the same direction.

Correlations were first computed within groups and then averaged, making use of the r-to-z transformation and weighting for sample sizes as recommended by Edwards (1950). This method avoids certain distortions that might arise if correlations were computed over the entire group of students (see, for example, Lindquist, 1940). In general, group refers to one experimental section or one control section; if there was more than one section of a kind, the correlations were computed within the section. In a few cases, moreover, the size of the group warranted a division into male and female students, and where this was possible, correlations were computed for the two sexes separately. This procedure yielded 14 experimental and 14 control sub-groups for seven courses taught in the Fall of 1958-59. The seven courses were: Freshman English (see chapter 9), General Psychology - I (see chapter 15), General Psychology - II (see chapter 16), Physical Science (see chapter 10), Political Science (see chapter 13), Principles of Economics (see chapter 11), and Principles of Marketing (see chapter 12).

Results. Before discussing these results, one important qualification should be entered that applies to all of the following remarks. Because of the large number of cases involved, very small correlations are statistically significant, and some of these involve very interesting patterns of results. But it must be stressed that none of these correlations is large enough to provide any stable basis for prescribing educational methods for individual students or types of student. Correlations of these low magnitudes may provide some additional basis for informed speculation about the process of education and its possible modifications, nothing more.

This need for caution is underlined when one considers the variability of the correlations. For example, the mean correlation between Part I of the finals and Vocabulary is the highest to be found in Table 2, yet the range of values of r is from .06 to .61. The difference in mean correlation between Part I of the finals and Academic Independence, .15 in the experimental group and -.08 in the control group, is significant ($p < .01$) and quite interesting. But the experimental group mean correlation of .15 involves values of r ranging from -.35 to .51. While the mean of .15 is significant and involves 4 significant positive correlations, it must be recognized that an investigator might conceivably have done only the one experiment involving a significant negative correlation, leading him to different conclusions from the ones suggested here. These fluctuations are not rare: Holland (1959), for example, studying the correlation of personality variables with college grades, found some correlations of opposite sign at institutions as similar as Harvard and Yale, or Radcliffe and Wellesley.

With due regard for these cautions we may proceed to examine Table 2.

Table 2

Intercorrelations Among Parts of Final Examination and Other Variables for Seven Courses: 1958-59

Experimental N = 380
Control N = 455

		1	2	3	4	5	6	7	8
								Educational Values	
		Final I	Examination II	III	IV	Vocabulary	Course Evaluation	Academic Independence	Spoon-feeding
1	Experimental		.36	.31	.31	.37	.17	.15	-.23
	Control		.36	.34	.25	.40	.20	-.08	-.04
2	Experimental			.26	.25	.24	.17	.04	-.09
	Control			.30	.13	.29	.12	-.07	-.01
3	Experimental				.24	.26	.06	.11	-.14
	Control				.29	.25	.17	-.14	00
4	Experimental					.24	.17	-.02	.02
	Control					.23	.10	00	.02
5	Experimental						.14	.08	-.10
	Control						-.02	-.03	-.05
6	Experimental							.09	-.24
	Control							-.13	.01

Note: for experimental group, if $r \geq .11$ $p < .05$, if $r \geq .14$ $p < .01$
" control " , if $r \geq .09$ $p < .05$, if $r \geq .13$ $p < .01$

a. Intelligence and self-directed study. The results give no support to the hypothesis that the more able students will adapt more successfully to self-directed study. There is a consistently positive, significant correlation of all parts of the final with vocabulary ($p < .01$), but the correlations are about the same for experimental and control groups. The mean correlation of the Dressel test of critical thinking, which is omitted from this table, with Part IV of the final was .02 for the experimental group, .18 ($p < .01$) for the control group, with a similar but less striking difference for its correlation with Part III; these differences are opposed to the hypothesis that high ability as measured by such tests predicts successful adaptation to self-directed study. This one very limited result suggests, rather, that self-directed study may have the effect of eliminating the correlation of performance with tests of intellectual ability, perhaps by giving students who need it an opportunity to learn to work independently, as required in Part IV of the final. In this connection, it should be recalled that in the Physical Optics course (Chapter 7) the students of lower ability, as tested by the AGCT, did better in the experimental group, while students of very high ability tended to do better in the control group.

Before leaving the question of ability, it is worth pointing out that the correlation of vocabulary with course evaluation is significantly positive ($p < .01$) in the experimental group, but near zero in the control group.

b. Academic independence and self-directed study. The correlations of performance on the final examination with the value placed on academic independence (see column 7 in Table 2) tend to be positive in the experimental group and negative in the control group, particularly for Parts I and III of the finals. A pattern having the same meaning occurs in the correlation of the finals with the spoonfeeding scores (see column 8 of Table 2): negative correlations of spoonfeeding with performance in the experimental group, coupled with zero correlations in the control group. This result gives some support to the hypothesis that students performed better on the final if the educational methods to which they were exposed conformed with their beliefs about higher education. Another interpretation of this result, however, would be that those students who had been able by the end of the semester to alter their educational philosophy to conform with the idea of self-directed study were also able to adapt to self-directed study in the sense of studying well under those conditions.

Finally, the correlations between course evaluation and attitude toward educational method should be noted. In the experimental group, course evaluation correlated positively with academic independence and negatively with spoonfeeding scores; in the control group it correlated negatively with academic independence; the differences between experimental and control groups are highly significant. In short, at the end of the semester, students tended to evaluate a course favorably if it conformed with their beliefs about educational methods.

Year in school, ability, and self-directed study

In this section will be presented the results of a group of analyses of variance involving 292 students in a course in general psychology. The course and the approach to the analyses of variance have already been described in chapter 15. In that chapter analyses were presented involving performance on four parts of the final examination as the dependent variables. In the present chapter the dependent variables used in the analyses were the measures of course evaluation and of attitudes toward educational method described above; to these four variables one has been added representing the combination of the academic independence and spoonfeeding scores for each student. Since these scores have opposite meaning spoonfeeding was subtracted from academic independence to give a combined score reflecting the student's total emphasis on methods favorable to self-directed study. Considering that the two scores used only seven of the eleven items in the hypothetical time inventory from which they were drawn, they are not merely complements of each other, and it was felt that their combination might give a stabler score than either one alone.

The independent variables in the analyses presented in this chapter are the same as before (see Chapter 15): treatment (self-directed study and control), ability (5 levels), sex, and year in school (2 levels, Freshman and Sophomore plus). The same method of repeated analyses was used.

Results. The statistically significant results for these analyses are summarized in Table 3. In this table only significant effects have been entered. The F-ratios associated with these effects are shown in Table 4.

Table 3

Summary of Analyses of Variance of Educational Values and

Course Evaluation in General Psychology

N = 292

Source of Variation	Acad. Ind.	Spoon Feed.	A.I.-S.P.	Absolute Evaluation	Relative Evaluation
Treatment	SDS Higher	-----	SDS Higher	Control Higher	Control Higher
Ability	-----	-----	-----	-----	-----
Sex	Females Higher	-----	-----	-----	-----
Year in School	Freshmen Higher	Sophomores Higher	Freshmen Higher	-----	Freshmen Higher

(Table 3 continued on next page)

Table 3 (continued)

Source of Variation	Acad. Ind.	Spoon Feed.	A.I.-S.P.	Absolute Evaluation	Relative Evaluation
Treatment x Ability	-----	-----	-----	-----	-----
Treatment x Year	-----	-----	Freshmen SDS higher than rest	-----	Freshmen Control higher than rest

Note: The students were divided into 5 ability levels, level 5 being the highest 20 per cent, etc.

As reported earlier in this chapter for seven courses in the 1958-1959 study, the experimental group was significantly higher on academic independence. For this course taken alone, the difference on the spoonfeeding variable was not significant, although it was in the same direction as for all seven courses. The method for deciding that this difference was insignificant in spite of the appearance of one relevant significant F-ratio has been described in chapter 15. For the combination of academic independence and spoonfeeding the difference was statistically significant, the experimental group being more favorably disposed than the control group toward academic behaviors associated with self-directed study.

There were no significant effects involving ability or the interaction between treatment and ability. As shown in Table 2, the difference between control and experimental groups in the correlation of ability and course evaluation suggested that there might be an interaction of treatment and ability in their effects on course evaluation, but this did not occur for this course.

Perhaps the most interesting results of these analyses are those concerning the effects of year in school. On all three variables involving educational values, Freshmen were more favorable to academic behaviors associated with self-directed study. There is also a significant interaction between treatment and year in school in their effects on the combined educational values variables. Inspection of the relevant means, shown in Table 5, makes it clear that the educational values of Freshmen and Sophomores are almost identical in the control group. It is the Freshmen in the experimental group whose educational values are significantly more favorable to academic independence than the others. This may be a useful result, since it suggests that Freshmen are open to change in their views about education.

As can be seen in the lower right hand corner of Table 3, there is a significant interaction between treatment and year in school in their effects on course evaluation. Freshmen in the Control group gave significantly more

Table 4

F-ratios Obtained in Analyses of Variance

N = 292

			Educational Values			Course Evaluation	
Source of Variation	df	Analysis	Academic Independence	Spoonfeeding	Acad. Ind. minus Spoonfeed.	Absolute	Relative
Treatment	1	TAY	13.20***	5.23*	12.35***	33.18***	35.27***
		TSY	5.7 *	.48	3.63	20.92***	20.18***
		TSA	8.59**	.81	5.67*	22.47***	19.28***
Ability	4	TAY	.70	.92	.66	.92	.23
		TSY					
		TSA	.25	1.77	.79	.64	.12
Sex	1	TAY					
		TSY	3.04	2.14	.17	.79	.18
		TSA	4.30*	2.42	.40	.18	.00
Year in School	1	TAY	6.09*	4.46*	7.03**	4.03*	7.24**
		TSY	3.17	7.14*	6.09*	3.29	4.55*
		TSA					
Treatment x Ability	4	TAY	.70	.46	.24	.17	.11
		TSY					
		TSA	.41	.16	.11	1.10	.49
Treatment x Year	1	TAY	.86	9.54**	4.38*	2.60	5.97*
		TSY	.13	1.43	.73	1.97	3.27
		TSA					

Note: TAY refers to analysis involving treatment, ability, year in school.
TSY refers to analysis involving treatment, sex, year in school.
TSA refers to analysis involving treatment, sex, ability.

* Significant at .05 level of confidence
** Significant at .01 level of confidence
*** Significant at .001 level of confidence

favorable evaluations than other students. Again, inspection of the means shows that Freshmen and Sophomores in the experimental group were about equally unfavorable toward the course. The Sophomore control students were neutral toward the course, their mean relative evaluation being exactly 8, midway between the Freshman controls (M = 10.2) and the SDS group combined (M = 6.6). The means and standard deviations for these results are shown in

Table 5, while the corresponding F-ratios are shown in Table 4.

Table 5

End-of-Term Course Evaluations and Attitudes Toward Independent Intellectual Work in General Psychology

		Relative Course Evaluation		Academic Independence minus Spoonfeeding[1]	
	N	M	SD	M	SD
Experimental					
Freshmen	47	6.7	3.4	51.4	5.3
Sophomores	108	6.6	3.2	47.8	6.6
Control					
Freshmen	38	10.2	3.0	47.6	4.7
Sophomores	99	8.0	2.9	46.8	4.9

[1] A constant of 50 was added to each score to eliminate negative numbers

Summary

Students in seven self-directed study courses evaluated these courses less favorably than did students in control courses. In spite of this unfavorable reaction, the SDS group seems to have moved toward a more favorable general attitude toward independent intellectual work. Even when the SDS group as a whole did better than the Control group in a particular course, their evaluation of that course was less favorable. In one course where the results could be examined in more detail, it appeared that there was an interaction between year in school and educational method, in the sense that the change toward a more favorable attitude toward independent intellectual work was confined to Freshmen in the experimental group.

The hypothesis that self-directed study would be more appropriate for the abler students than for the less able ones was not borne out: although there was a low positive correlation between ability measures and final examination performance, this applied equally to experimental and control groups.

In the experimental group there was a low positive correlation between the student's final examination performance and the extent to which he favored doing independent intellectual work; in the control group this correlation was negative. This pattern gives some support to the hypothesis that students do better academically if the educational methods they are subjected to conform with their own views about such methods. Similarly, there was a low positive correlation between course evaluation and academic independence in the experimental group and a low negative correlation in the control group, suggesting that students give more favorable evaluations of courses taught in a manner conforming with their own views about education.

None of these results are considered to provide the basis for a direct prescription for different educational methods to be used with different students. Rather, they may shed a little additional light on the complex relationships among abilities, attitudes, and performance in different educational settings.

Chapter 19

Effects of Self-Directed Study on Thinking During Lectures

There is a gulf between the lecturer and his audience, the more he talks the wider it grows and the less he knows about them. What do college students think during lectures? The lecturer does not know. There is at least some evidence that students are not at their peak of intellectual engagement during lectures. Brinkley asked students to rank a variety of learning situations as to the degree of mental activity elicited by them. Mental activity was defined as being "concentrated on the subject under consideration, in an attitude of active thinking...the opposite of mind wandering, mental laziness, passive receptivity" (1953, p. 536). The lecture situation was ranked seventh out of ten, the entire list, in order of the mental alertness they reportedly elicit being as follows: written examination, group discussion, preparing a paper, studying alone, oral quiz recitation, extra-curricular activity, science laboratory, lecture, studying with others, listening to a student report.

In a pioneering effort, Bloom (1953) studied students' thought-processes during lectures and discussions. He used a method of "stimulated recall" in which students were interviewed individually while listening to a tape-recording of a class meeting they had attended a few days before. From content analyses of the interview protocols Bloom found that lectures evoke thinking that is primarily suited to assimilating information, while discussions evoke more thinking that involves complex problem solving. Also, discussions evoked more thinking about the self, particularly about the relation of the self to the problem of mastery of the subject matter under discussion. In lectures, it is easier for the student to avoid active thought about the course, either by mind-wandering or simple passivity: "no one can threaten him during the lecture so long as he wears a mask of attention" (1953, 166).

Bloom's emphasis is largely on the difference between the immediate situation confronted by the student in the lecture and in the discussion. In addition, however, it is plausible that the effects of a specific situation, such as a lecture, might depend on variations in the total ecology of which the situation is a part. There are two reasons to expect that students in self-directed study courses may listen to lectures more effectively than students in conventional courses. First, since they attend fewer lectures they may simple become less jaded: a lecture once a week may be refreshing, three times a week may glut the market. Of course, it should be remembered that in the present experiments, the total number of lectures the student attended in a week was only reduced from about 15 to about 13, since only one of his courses was affected by the experiment. Secondly, if the student becomes more actively engaged with the material of the course, as required of him, one hopes, by the reduction in the number of lectures, he may be a more active listener because he has learned how to grapple with the subject matter at hand.

In order to explore this approach more fully, it was decided to capitalize on the fact that in the Fall of 1958 we would have several experimental courses going on at the same time. A method of investigation was needed that would be less time-consuming and expensive than Bloom's. In a series of preliminary interviews, we found that students interviewed immediately after a class period could give a fairly full description of their thought-processes during the class, although even a one hour delay reduced the recall noticeably. Therefore, the method of stimulated recall could be dispensed with if we could find students who were free immediately after the experimental classes. In addition, it was important to eliminate the necessity for individual interviews and subsequent analysis of long interview protocols. To meet this need an inventory of thought-processes in the classroom was developed along the following lines: first the student is asked to write brief notes on what he had been thinking, feeling and doing during the lecture. The he is asked to rate his level of attention during the lecture (on a 5-point scale in the preliminary instrument, and a 6-point scale in the one used in the main study). Then he is asked to indicate the distribution of his efforts over a number of activities by indicating "the per cent of your total effort, energy, or attention that you spent on each activity during your last class." The activities listed were derived in part from Bloom's work. Finally, the student is asked to write down briefly any questions or problems that occurred to him during the lecture.

A pilot study conducted during the summer of 1958 tested the technique by means of an experiment in which some students were given brief training in critical thinking: as compared with a control group, the experimental group showed certain predicted changes in their thinking in a lecture following the training period. On the basis of this pilot study, the questionnaire was revised slightly, primarily by shortening it, for use in the self-directed study program.

Procedure

Eight lectures were used as field situations for the study of the stream of thought in the classroom; four lectures were in self-directed study classes while the other four were in the corresponding control classes. The courses involved were Physical Science, Political Science, Principles of Economics, and General Psychology - II. A roster was compiled of students who had no class immediately following these selected lectures. Each of these students was telephoned and asked to participate in an experiment. About 80 per cent of those contacted participated. Immediately after the lecture the participants came to a room near the classroom they were coming from and were given the questionnaire described above. Neither the students nor the lecturer were given any advance indication that an experiment of this kind was occurring.

It should be mentioned that a fifth course, General Psychology - I, was originally involved. One of the hazards of not informing the teachers of our plans took its toll: the teacher of the experimental section gave out a questionnaire which took the first fifteen minutes of the class, and the effects of which noticeably colored the students' thinking during the remainder of the hour. It was therefore decided to drop this group and the corresponding control group from the results.

Results and Discussion

There was a distinct difference between experimental and control groups in the students estimates of their level of attention. Shown in Table 1 is the exact wording of the items in the question, which asked the students to "choose the sentence that best describes the level of your attention in your last class." The level of attention was higher in the experimental group. Dividing the respondents into those who "kept the thread" and those who "lost the thread" (items a, b, c versus items d, e, f), and forming a 2 x 2 contingency table, the difference is shown to be statistically significant (chi-square = 3.86, $p < .05$).

Table 1

Level of Attention Reported by Students Attending

Lectures in Experimental Program

Item	Number reporting Experimental N = 27	 Control N = 27
a. My mind never wandered	1	7
b. My mind wandered a little, but I did not lose the thread	18	6
c. My mind wandered quite a bit, but I did not lose the thread	5	5
d. I lost the thread occasionally	1	4
e. I stopped following but I could probably remember most of it because I was listening with one ear	1	2
f. My mind was on other things most of the time	1*	3

*This student reported that she had spent the whole hour studying for another course.

It will be noted that in spite of the generally higher level of attention in the experimental group there was only one student in that group who reported that his "mind never wandered," while there were seven in the control group who gave that response. Examination of the protocols clarifies what happened. Five of those seven individuals were in one course. The instructor of that course made a practice of calling on students to answer questions in a way that seems to have generated almost universal anxiety. Some typical remarks follow, unlike any made by students who had attended other lectures: "Quite often thoroughly confused; nervous, afraid that if he did ask me a question that I could not answer properly and would get cut down." "Unable to answer question addressed to me. Felt ridiculous." "...he called on me. I thought now I'm in for it, I haven't read the material. But from his previous words, I remembered and answered correctly, but where that answer came from I don't know--lucky guess......I left the lecture determined to read the assignments over Christmas." "Dreaded the thought of being called on in class and was relieved when he called on others. Became angry when he continued to lecture at such a rapid pace. Decided to study...this weekend after threat from Mr. --- to class in general...Had the feeling at the end of class that Mr. --- was angry too." Frequent mentions were made of this lecturer's rapid pace, of feelings of confusion, of renewed determination to study. It is clear that this teacher was not only feared, but respected for his ability and that he succeeded in riveting the students' attention on the lecture.

Whether this is an unmixed blessing is another question. On the next part of the questionnaire, the respondents were asked to estimate "the per cent of your total effort, energy, or attention that you spent on each activity during your last class." A list of eleven activities was provided, as shown in Table 2, together with the results. The results for the control group class in which so much anxiety was generated indicate that these students devoted considerably more effort than any other class to items 2 and 3, having to do with trying to take in the material, but they devoted less effort than any other class to the kind of critical thinking reflected in items 5 and 6.

Table 2

Mean Per cent of Effort Reported Spent on Various

Cognitive Activities While Listening to a Lecture

	Per cent Effort	
	Experimental	Control
Cognitive Activity	N = 27	N = 27
1. Listening to what was being said without thinking about it very much	8.3	6.4
2. Taking notes	25.6	26.1
3. Following what was being said	24.8	29.6
4. Relating what was being said to other information on the same subject (previous classes, your own experience, etc.)	6.8	7.6
5. Evaluating what was being said as to its truth, plausibility, logical coherence	8.7	6.3
6. Evaluating what was being said as to its meaning and value for you	5.8	2.6

(Table 2 continued on next page)

Table 2 (continued)

Cognitive Activity	Per cent Effort Experimental N = 27	Per cent Effort Control N = 27
7. Thinking about your own adequacy with respect to the subject matter or in comparison with other students	4.5	3.2
8. Thinking about personal characteristics of the teacher and/or other students in the class	5.0	4.1
9. Incidental thoughts about words and phrases used in the class	1.4	3.0
10. Thinking about persons, objects, or events _not_ in the classroom and not related to the subject	5.1	8.4
11. Miscellaneous activities such as looking around the room or out the window, listening to random noises, or attending to your own bodily sensations	3.4	3.1

Note: the questionnaire also had a space for "other" activities and requested the respondent to explain any entries in that space. Very few students used it, and all but one of these entries could be classified as representing one of the above items. Omitted from the table is the response given by one student that she had spent 75 per cent of her effort studying for another course.

For a number of reasons tests of significance have not been attempted for the differences between experimental and control group means shown in Table 2, but chiefly because each mean is based on a small number of cases (from 3 to 11) from four different lectures, and there is no particular reason to believe that all experimental or all control lectures are enough alike to produce clearcut similarities in the fine details of student reactions to them. If enough data could have been collected, it would have been possible to find a satisfactory statistical approach to the variation between lectures within educational treatments.

The items in Table 2 have been compressed in Table 3 into four major categories. Items 1, 2, and 3 may be regarded as relatively passive efforts to record and assimilate the lecture. Items 4, 5, and 6 have been grouped as representing efforts to think actively about the material of the lecture. Item 7, dealing with thoughts about personal adequacy provoked by the lecture, has been kept by itself. Items 8-11 have been grouped as representing varieties of mind-wandering.

Table 3

Mean Per cent Effort Reported Spent on Four Kinds of Cognitive Activity While Listening to a Lecture

	Per cent Effort	
Kind of Cognitive Activity	Experimental N = 27	Control N = 27
Passive assimilation of content of lecture (items 1, 2, 3)*	58.7	62.1
Active thinking about content of lecture (items 4, 5, 6)	21.3	16.5
Thinking about personal adequacy with regard to subject matter of course (item 7)	4.5	3.2
Mind-wandering (items 8, 9, 10, 11)	14.9	18.6

*For exact wording of items see Table 2

The striking thing about Table 3 is its similarity to Bloom's results (1953). In his comparison of thought-processes during lectures and discussions, he found that the latter evoked less passive assimilating, more active thinking, more concern about personal adequacy, and less mind-wandering. These are exactly the differences between the response to lectures of the self-directed study group and the control group. Due to differences in method, there is no direct way of comparing the magnitude of the results of the two studies, but Bloom's results probably involve larger differences.

The sheer variety of cognitive activities engaged in is itself of some interest, although its meaning may be ambiguous. On the one hand, a richly varied cognitive life may indicate versatility and freedom of cognitive movement; on the other hand, it may merely indicate diffuseness of thought. In the present study, it was possible to get some idea of the variety of each individual's thought-processes simply by counting the number of categories he used in describing his thinking during the lecture. Two individuals who gave very different percentage distributions among the categories shown in Table 2 might still make use of the same number of categories.

It was found that for the total group of 11 categories, the experimental group showed a greater variety of thought, (chi-square = 3.71, $p < .10$) but the difference was not quite significant. For the first five categories in Table 2, those directly involving thinking about the lecture, the difference was significant (chi-square = 4.07, $p < .05$). It is interesting that the experimental group reported a somewhat greater variety of types of mind-wandering but, in terms of reported percent effort, did less of it altogether.

Considering the fact that the experimental students reported a higher level of attention and greater effort expended on thinking about the lecture, the greater variety of thought they displayed can probably be ascribed to greater freedom of cognitive movement resulting from self-directed study. In a strong form, this hypothesis might mean that the experimental group were more deeply and personally engaged with the material of the course, less in the lecturer's thrall. In a weaker form, it might merely mean that as listeners they were fresher because they were less jaded with lectures.

With more data, other interesting analyses would become possible. For example, there is a suggestion that students who "lost the thread" of the lecture were among those who spent little effort on taking notes, although there were also some who reported a high level of attention but not much effort expended on note-taking. Moreover, those who devoted considerable effort to evaluating the lecture (items 5 and 6) tended to spend little effort on note-taking. Before results of this kind could be interpretable it would be necessary to know something about the relation between effort and skill in note-taking. Some students may simply not bother to take many notes, others may take notes effortlessly, and others may do it at the expense of thinking about the material they are hearing. Although there have been a few studies in this area (see, for example, Eisner and Rohde, 1959) not enough is known to settle such questions.

One last result remains to be reported. After the respondents had filled out the parts of the questionnaire giving quantitative estimates of their thinking during the lecture they were asked to "list...all questions, problems, issues or doubts related to the subject of the course that you thought of during your last class. Write them all down regardless of whether they just flashed through your mind or you spent some time thinking about them."

The students' responses to this question were content analyzed by the method of blind analysis into the following four categories: questions regarding the subject matter of the course, questions regarding administration of the course (tests, assignments, etc.), questions regarding the relation of the self to the course, and questions unrelated to the course. A reliability check by another rater on a sample of 20 protocols indicated very high agreement, so only the ratings of the first rater were used for the results. Considering all kinds of questions combined, the median number of questions raised by the experimental group was 5, while the control group median was only 3. Since the main difference between the two distributions was at the upper extreme of question-raising behavior, a 2 x 3 contingency table was used, and the difference between the groups was significant (chi-square = 8.56, df = 2, $p < .02$). Most of the questions thought about in both groups had to do with the subject matter of the course; for this category the difference was again significant (chi-square = 4.07, df = 1, $p < .05$). This result gains in interest when it is remembered that in another study, self-directed study groups tested one year after a course had ended raised more questions than control students about the subject matter of the course (see chapters 2 and 3).

Summary

Students were given a questionnaire immediately after they had attended one of eight lectures; four of the lectures were in self-directed study courses where students were attending only one lecture per week, while the other four were in control courses where students were attending three lectures per week. The self-directed study group exhibited a higher level of attention, more effort expended on critical thinking, and more effort expended on thinking about the relation between the self and the subject matter of the course. During the lectures, the experimental group also thought of more questions about the subject matter of the course. There is no basis in the results for deciding whether these differences in patterns of cognitive activity while attending a lecture were due directly to the reduction in number of lectures attended and consequent freshness of the listeners, or whether it was due to greater personal involvement of the self-directed study group with the subject matter of the course. Both factors might be operating.

Part IV

The University and Self-Directed Study

Every educational experiment is introduced into some particular setting, a context which gives the experiment its distinct significance. A variation in teaching methods which is trivial in one setting may be substantial in another. For example, the introduction of small group discussions might be pointless in an academic climate where coffee-houses near the University were constantly filled with students passionately discussing their work; but if the students have no tradition of this kind it may well behoove the teacher to provoke such discussions in his classroom.

The following chapters are three approaches to the problem of describing the institutional setting in which the research program evolved. First, the results are presented of a comparison of the University with other institutions, using the College Characteristics Index, an instrument for which information is available about many institutions of higher education. Secondly, an examination is made of how students spend their time at the University. Finally, the goals of education, as viewed by students who had participated in the research program are examined.

Throughout this section, two questions remain uppermost. To what extent does the University ordinarily provide the student with the occasion to do independent intellectual work and to increase his capacity to do it? How much of a departure from prevailing practices and ideas did the Self-Directed Study Program actually represent?

Chapter 20

The University As An Environment

The College Characteristics Index (CCI)

The CCI was developed by Pace and Stern as a method of systematically characterizing institutions of higher education. Based on the Murray system of environmental press (Murray, 1938), the CCI provides scores for an institution on each of 30 scales. As the instrument has been administered in many institutions, nationwide norms are available. In one study, for example, norms are reported for 32 schools, subdivided in 6 categories. One of these categories, "universities," is the most logical basis for comparison with the University of Colorado. The normative data for this category is based on 543 students distributed over 7 universities, both public and private. Considering the nature of the test, this is probably an adequate number of individuals to establish a norm. The respondent is not asked to describe himself: the questions on the CCI are all formulated so that they relate to practices prevalent at the school under consideration. Thus, although 50 or 100 students per institution might be a small sample for some purposes, the sample size for the CCI ought to be compared with the small number of informants used in anthropological field work, rather than with the large samples used in opinion polling (Stern, 1958).

In the present study, the CCI results will be analyzed first according to the general approach followed by Pace and Stern, using the norms they obtained. In addition to the comparison of results for the University of Colorado with the Universities' norm, the College of Arts and Sciences will be compared with the norm for 7 Liberal Arts Colleges. The latter comparison provides some basis for understanding the difference between introducing self-directed study at a large university and introducing it at a liberal arts college.

In the light of Thistlethwaite's recent critical analysis of the CCI the results will also be examined according to his approach. The latter provides a basis for comparing the University with institutions rated highly productive or highly unproductive, in the sense that they turn out many or few graduates who go on to attain the Ph.D. degree.

The Sample. During the fall of 1960, 600 letters were sent to a sample of students at the University of Colorado, stratified by year in school and sex. The students were invited to participate in a study of the University of Colorado, requiring about two hours of their time. Of those invited, 188 appeared, so that the sample actually obtained is not necessarily representative in all respects. For the purposes of the present study, however, this is not necessarily a serious failing. Pace and Stern stress the need for a well-informed sample, rather than a representative one (Stern, 1958). To ensure meeting this need, only those Juniors and Seniors who were enrolled for a minimum of 12 semester hours (the official definition of a full-time student) and who had attended the University of Colorado for two or more years (i.e., excluding recent transfer students) were included. Although this reduced the sample to 88 students, it was felt that the application of these criteria would guarantee that the CCI results would be based on a panel of students well-informed about the University.

Results: CCI Scales

Pace and Stern (1960) have identified five major groups of scales, each of which refers to an important characteristic of the educational environment. In this section, each characteristic will be taken up in turn. The University of Colorado will be compared with normative data based on responses from 543 students at 7 universities. In addition, the College of Arts and Sciences at the University of Colorado will be compared with normative data based on 460 students at 7 non-sectarian liberal arts colleges.

In considering these results, the reader should keep in mind that they represent a description of the University as the students see it. Whether or not there is any other psychologically real university is a separate question.

Abstract thought, aesthetic interest, intellectual vitality and achievement. Of the 30 scales on the CCI, 7 scales form a group reflecting a broad environmental press toward an energetic conern for intellectual and aesthetic matters. Listed below is the name of each scale and a few of the items composing it. The word true or false in parentheses indicates the answer contributing to a high score on the scale.

Humanism: A lecture by an outstanding literary critic would be poorly attended (false). Humanities courses are often elected by students majoring in other areas (true). In many courses the broad social and historical setting of the material is not discussed (false).

Reflectiveness: There are many facilities and opportunities for individual creative activity (true). Modern art and music get little attention here (false).

Understanding: There is a lot of emphasis on preparing for graduate work (true). Most students have very little interest in round tables, panel meetings, or other formal discussions (false). "Alma Mater" seems to be more important than "subject matter" at this school (false). In class discussions, papers, and exams, the main emphasis is on breadth of understanding, perspective and critical judgment (true).

Achievement: Most courses require intensive study and preparation out of class (true). Students who work hard for high grades are likely to be regarded as odd (false). Learning what is in the text book is enough to pass most courses (false).

Sentience: To most students here art is something to be studied rather than felt (false). There is a lot of interest here in poetry, music, painting, sculpture, architecture, etc. (true). In papers and reports, vivid and novel expressions are usually criticized (false). There are paintings or statues of nudes on the campus (true).

Objectivity: No one needs to be afraid of expressing extreme or unpopular viewpoints at this school (false). A well reasoned report can rate an A grade here even though its viewpoint is opposed to the professor's (false).

Energy: Students get so absorbed in various activities that they often lose all sense of time or personal comfort (true). Class discussions are typically vigorous and intense (true). Many lectures are delivered in a monotone with little inflection or emphasis (false).

Table 1 shows the norms for 7 universities and for 7 liberal arts colleges, and the results for the University of Colorado and for the College of Arts and Sciences. It can be seen that the University is much like other universities in the country. Students in the College of Arts and Sciences, however, share in an educational environment very different from students in liberal arts colleges. The latter enjoy an environment with greater press toward intellectual and aesthetic pursuits, greater acceptance of academic ambition and intellectual vigor.

Table 1

College Characteristics Index: Standard Scores on Scales Reflecting Environmental Press Toward Abstract Thought, Aesthetic Interests, and Intellectual Vitality

	Univ. norm	Univ. Colo.	L.A. norm	A & S Univ. Colo.
N	543	88	460	60
Scale				
Humanism	4.69	4.98	6.61	5.22
Reflectiveness	4.48	5.67	6.54	5.91
Sentience	4.13	5.00	6.69	5.14
Understanding	4.37	4.27	6.32	4.20
Objectivity	4.49	4.22	5.96	4.21
Energy	4.06	4.52	6.25	4.54
Achievement	4.19	3.95	5.96	3.85

Note: The Universities norm is based on 7 institutions; the Liberal Arts Colleges norm is based on 7 institutions (Pace and Stern, 1960).

The "Fun culture". Two scales form a group reflecting this characteristic, the play scale and the sex scale. The latter is actually a group of items referring largely to the institution of dating. As above, a few sample items for each scale are given, and the parentheses after each item indicates the answer contributing to a high score on the scale.

Play: Students really get excited at an athletic contest (true). There are lots of dances, parties, and social activities (true). There is very little studying here over the week-ends (true). Everyone has a lot of fun at this school (true). It's easy to get a group together for card games, singing, going to the movies, etc. (true). Students are very serious and purposeful about their work (false).

Sex: Students spend a lot of time talking about their boy or girl friends (true). Nearly everyone here has a date for the weekends (true).

Table 2 shows the norms for 7 universities and 7 liberal arts colleges, and the results for the University of Colorado and for the College of Arts and Sciences. It can be seen that the universities are quite different from the liberal arts colleges, the former providing much greater environmental press toward participation in the fun culture. The University of Colorado is above the norm for universities, but the difference is not statistically significant. The College of Arts and Sciences is significantly above the norm for liberal arts colleges in its emphasis on the fun culture.

Table 2

College Characteristics Index: Standard Scores on

Scales Reflecting Environmental Press Toward

Participation in the Fun Culture

	Univ. norm	Univ. Colo.	L.A. norm	A & S Univ. Colo.
N	543	88	460	60
Scale				
Play	6.14	6.70	3.27	5.36
Sex	6.30	7.10	4.06	7.21

Note: The Universities norm is based on 7 institutions; the Liberal Arts Colleges norm is based on 7 institutions (Pace and Stern, 1960).

The other groups of scales were the following: a group reflecting environmental press toward intellectual competitiveness; a group reflecting press toward concern for human relations; and a group reflecting press toward adolescent rebelliousness. On none of these was the University of Colorado or the College of Arts and Sciences consistently and significantly different from its respective normative group.

The Academic Environment and Student Achievement

In the previous section the University of Colorado was compared with norms for a group of other institutions. Such a comparison reveals whether or not the University is unusual, but it does not provide a basis for assessing the value of the university environment. In order to deal with questions of this kind, Thistlethwaite used the CCI to compare the academic environment provided by "productive" institutions with that provided by "unproductive" institutions. His index of productivity was "the discrepancy between a college's expected rate of producing ... Ph.D.s (as predicted from its enrollment of talented students) and its actual rate of productivity" (1959, p. 184). Separate indices were computed: one for natural science Ph.D.s and one for Ph.D.s in the arts, humanities and social sciences. Thistlethwaite found that the college environment had a distinct effect on productivity, and that the environment conducive to high productivity in natural science was different from the environment conducive to high productivity in the scholarly fields.

Although Thistlethwaite found that the CCI scales as developed by Pace and Stern were useful in the analysis described above, he was able to improve on them by means of an item analysis. In the present study, the responses of students at the University of Colorado are compared with responses of students at the most productive and least productive institutions. Table 3 shows this comparison for items discriminating productivity in the natural sciences; Table 4 shows this comparison for items discriminating productivity in the arts, humanities and social sciences.

In the last column of each of these tables a minus sign indicates that on this item the University of Colorado more closely resembles the least productive institutions and a plus sign indicates closer resemblance to the most productive ones. It is clear the University of Colorado resembles the least productive institutions more closely than it does the most productive institutions. This is true for both indices of productivity. Students at more productive institutions receive a greater intellectual challenge, benefit from a better library, and are less closely supervised than at the University of Colorado. Items on which there are the sharpest differences refer to emphasis on preparation for graduate work and the extent to which "professors really push the students' capabilities to the limit."

In a number of ways, the least productive institutions are characterized by the extent to which the student is not given responsibility for his own education: assigned seats, frequent tests, few elective courses, little library work, and much supervision of student activities--these are attributes of the least productive schools. While not extreme in these respects, the University of Colorado is somewhat more similar to the least productive institutions than to the most productive ones.

Discussion

It is clear from the above results that the University of Colorado provides a distinctive environment. But it should be borne in mind that the characteristics of this environment derive not only from the faculty, administration, and physical setting: they derive also from the students themselves. In Thistlethwaite's studies an attempt was made to control for differences in intellectual ability of students entering the different institutions studied. But recently Astin (1961)

Table 3
Comparison of University of Colorado with Institutions Most and Least Productive of Ph.Ds in Natural Sciences[1]

Item	Percent of "Yes" Responses Most Prod.	Least Prod.	Univ. Colo.	Sign[2]
In talking with students faculty members often refer to their colleagues by their first names.	34	22	24	-
Faculty advisors or counselors are pretty practical and efficient in the way they dispatch their business	64	71	41	+
Personality, pull, and bluff get students through many courses.	16	34	51	-
The professors really push the students' capabilities to the limit.	60	45	11	-
Course offerings and faculty in the natural sciences are outstanding.	87	75	55	-
The school is outstanding for the emphasis and support it gives to pure scholarship and basic research.	84	77	49	-
In many classes students have an assigned seat.	35	62	74	-
Student organizations are closely supervised to guard against mistakes.	18	30	35	-
Student publications never lampoon dignified people or institutions.	40	12	7	-
It is easy to take clear notes in most courses.	75	84	47	+

Footnotes for Tables 3 and 4

[1] "Most productive" refers to 12 of 36 institutions (N = 916) which produced the largest number of undergraduates obtaining Ph.Ds. The converse is true for "least productive" (Thistlethwaite, 1959).

[2] See text for explanation.

[3] This item is only in essence the same as the item used by Thistlethwaite. In form 457 of the CCI it appears as: "Most courses require a lot of library work." University of Colorado data were obtained with form 1158 where the item appears as above. Other items are identical for both CCI forms.

Table 4
Comparison of University of Colorado With Institutions Most and Least Productive of [illegible]
Arts, Humanities, and Social Sciences[1]

	Percent of "Y[illegible]es			
	Most Prod.	L[illegible] Pro[illegible]	[illegible]	Sign[2]
Course offerings and faculty in the social sciences are outstanding.	79	2[illegible]	48	-
The library is exceptionally well equipped with journals, periodicals, and books in the social sciences.	96	61	58	-
Professors often try to provoke arguments in class, the livelier the better.	53	34	28	-
Faculty members put a lot of energy and enthusiasm into their teaching.	87	72	57	-
Many courses stress the speculative or abstract rather than the concrete and tangible.	88	56	73	+
There is a lot of emphasis on preparing for graduate work.	81	52	20	-
Students almost never see the professor except in class.	29	48	43	-
Faculty members are impatient with students who interrupt their work.	14	23	22	-
If a student fails a course he can usually substitute another one for it rather than take it over.	70	12	39	-
Students are expected to work out the details of their own program in their own way.	72	52	65	+
Most courses require intensive study and preparation out of class.[3]	65	15	49	+
In class discussions, papers, and exams the main emphasis is on breadth of understanding, perspective, and critical judgment.	95	80	69	-
The school offers many opportunities for students to understand and criticize important works in art, music, and drama.	90	43	60	-
Frequent tests are given in most courses.	26	64	45	-

(For footnotes see Table 3)

was able to show that after controlling for ability a large part of the remaining variation in college productivity was due to another characteristic of entering students: their academic aspiration level, i.e., the strength of their interest in graduate work.

In any event, whether in the past the college environment has made the student or the student has made the environment, it remains the task of the educator to understand that environment and to do what he can to shape it into an effective educational instrument.

With regard to self-directed study, it seems obvious that the University of Colorado differs from certain other institutions in important respects. The introduction of self-directed study at the University of Colorado may be a much greater departure from established practices than it would be at a small liberal arts college with an intensely intellectual atmosphere and considerable emphasis on independent work already prevailing.

Chapter 21

How Students Spend Their Time

One direct, useful way of describing an institutional environment is to examine what the inhabitants actually do. How do the days pass? What pursuits occupy the most time and the least time? Although the relationship cannot be a simple one, the time devoted to various pursuits must in some way reflect the goals of the institution, if not the stated goals then the ones actually advanced.

In some measure, the distribution of time reflects the individual student's own decisions, but in some measure it reflects the kind of requirements set for him by his teachers and by other features of the institutional environment. In any case, the pattern of human activity, however it has come about, is itself an essential feature of the human environment.

Probably the most thoroughgoing studies of this problem are the work of Barker and Wright (see for example, Barker & Wright, 1955). Their method, however, involving direct observation by a team of field workers, is extremely time-consuming and expensive. For the present study, what was needed was a simple, economical approach. An inventory was developed, using the students' self-reports of the distribution of time over various activities. While this method may have certain disadvantages, for the present purposes it also had its merits. At some of the most interesting points, it is difficult to find out what the individual is doing without asking him. For example, if a student is sitting quietly in a classroom, he may be listening attentively, thinking critically, or daydreaming. There is reason to believe (Bloom, 1953) that a good way to find out what he is doing is to ask him.

In developing a questionnaire on this subject, one of the great difficulties is the complexity of the question,"How do you spend your time?" At any given moment, an individual is doing a very large number of things, any or all of which might be included in a description of what he is doing. A full description of a student's week could be of any length, limited only by the endurance of the writer.

The approach taken here to this dilemma has been to forego the aim of a complete description in which one tries to say everything that might be said, and to substitute the aim of an exhaustive description in which one tries to say something relevant about every moment in the period of time under consideration. This approach is analogous to methods used in genetics; Mendel, for example, made no effort to give a full description of every sweet pea, but he was careful to examine each specimen (i.e., exhaust the set of specimens) with regard to certain relevant characteristics. This decision to be exhaustive but not complete is one way of simplifying the description of the complex entity, how students spend their time.

A further way of simplifying the description of such a complex entity is to divide it into parts, and then to subdivide each part, continuing this process of subdivision as long as meaningful descriptive terms for the subparts can be found. A crude form of this method has been previously used, with some success in describing educational programs (Gruber, 1961).

It follows from information theory that, in the ideal case, such a description will be maximally efficient if every subdivision divides the given entity into exactly two equal subparts (Attneave, 1959). As applied to the description of the behavior occurring in a period of time, such as a week, the product of this process of successive subdivision can be referred to as a behavior tree. Although the reader will see below that it was not possible to use a tree with equal bifurcations at every branching, this was the guiding idea behind the development of the inventory. It was a helpful idea, both in guiding the development of the inventory and in providing the students with a format in which to give a coherent answer to a complex question.

The Inventory: How Students Spend Their Time

Preliminary interviews. In a series of preliminary interviews, 10 students were asked to describe the distribution of their time over activities. First, the interviewee was asked to divide his time in a typical week in the academic year into two equal or nearly equal categories. Then he was asked to subdivide each of these categories into two equal or nearly equal categories, continuing the process of subdivision as long as it seemed meaningful to the interviewee. As he worked, he made a branching diagram of the process of subdivision. Once he had developed the "tree" of his week's activities, he was asked to go over the categories and indicate the number of hours devoted to each activity. The interviews lasted about an hour each, and although the task required considerable thought it seemed to be a fairly natural way to arrive at a systematic description of the allocation of time. If at any branching point the interviewee found that the restriction to exactly two branches was unnatural, he was permitted to use more than two, but was encouraged to keep the number small.

On the basis of the behavior trees produced in these interviews, an attempt was made to develop a modal tree, that is, a set of behavior categories, arranged in a branching hierarchy that would provide a reasonable basis for all students to use in describing how they spend their time. Thus, in the instrument as finally developed, the behavior tree, with labelled categories, is provided; the student's task is only to fill the blanks indicating the amount of time he spends on each kind of behavior.

The modal tree. In the form used in the main study, the student was asked to subdivide 100 hours in a typical waking week into time spent in three major categories: academic work, nonacademic activities, and paid employment. Time spent in paid employment was not further subdivided. Time spent in nonacademic activities was further subdivided in some detail but the results will not be reported here. The time spent in academic work was further subdivided in three successive steps as shown in Table 1. The final subdivisions have

Table 1

How Students Spend Their Time

Academic work

- A. Time spent in class
  - 1. Listening to instructor
    - a) Trying to follow the instructor
    - b) Thinking actively and critically about the material being presented.
  - 2. Discussion and Recitation
    - c) Answering questions and doing exercises to show that you have done assignments
    - d) Real discussion that involves thinking actively and critically.
  - 3. Laboratory work (include field work)
    - e) Following instructions, with little or no opportunity to make decisions.
    - f) Carrying out assignments with some important decisions about what to do and how to do it left up to the student.
    - g) Independent work, with most of the important decisions left up to the student.
  - 4. Taking tests
    - h) Objective tests (true-false, multiple choice, completion, etc.)
    - i) Problems involving applications of methods applied by instructor or textbook.
    - j) Essays and problems in which major emphasis is on your own thought.
- B. Time spent out of class.
  - 1. Reading
    - k) Reading assigned textbooks.
    - l) Assigned reading other than textbooks.
    - m) Reading unassigned material connected with academic work.
  - 2. Discussing
    - n) Reviewing and clarifying notes, textbooks, classwork.
    - o) Thinking actively and critically about your academic work.
  - 3. Writing
    - p) Routine themes and reports.
    - q) Papers in which major emphasis is on your own thought and investigation.
  - 4. Perfecting skills
    - r) Doing problems
    - s) Practicing (for example, music)
    - t) Memorizing
  - 5. Research and creative work
    - u) Scientific (experimental, theoretical)
    - v) Aesthetic (creative writing, composing, artistic creation etc.)

been lettered continuously for ease of reference. It will be noted that these final subdivisions are formulated so as to make it possible to estimate the amount of time that the student's academic work can be said to be active, independent, and self-directed on the one hand, or passive, dependent, and teacher-directed on the other. For example, the student in the classroom listening to the instructor may be "a. Trying to follow the instructor" or "b. Thinking actively and critically about the material being presented."

The outline of the section of the modal tree shown in Table 1, is not set forth in the same format as in the actual questionnaire. The latter was arranged as a series of branching charts, with spaces provided for each student to insert the number of hours allotted to a particular activity in a typical week and in an ideal week.

The students to whom the modal tree was administered were the same as those who received the College Characteristics Index, described in the previous chapter. In the results that follow all students are included who were classified as full-time students (registered for 12 or more semester hours), regardless of how long they had been at this University. In addition, the modal tree was administered in two sections of Engineering English, a course taken by all Freshman Engineers, in order to increase the number of cases in that group.

Results

In presenting the results, the subjects are divided into 12 sub-groups according to year in school, sex, and college (Arts & Sciences or Engineering). Not enough data was obtained from students in other divisions of the University. Median results for the 12 subgroups are kept separate in the tables that follow. To save space, information regarding variability has been omitted from these tables but will be made available separately.

Table 2 shows the median allocation of time to the three categories at the first level: nonacademic activities, academic work, and paid employment. In the College of Arts and Sciences time is divided about equally between academic work and nonacademic activities, with slightly more time given to the latter. The discrepancy between actual and ideal allocations indicates a general desire to alter the balance in the direction of doing more academic work. Engineering students report longer hours actually devoted to academic work, and they too would like to devote more time to it than they do at present, although the discrepancies between actual and ideal are smaller than for Arts and Science students.

Table 3 shows the breakdown of time spent in the classroom into various categories. The various subgroups in Arts and Sciences spend between 16 and 20 hours per week in class. In every subgroup half or more tha half of this time is spent in listening to the instructor. Discrepancies between actual and ideal indicate a consistent desire for somewhat less listening and more discussing. Engineering students spend more time in class than do students in Arts and Sciences, the difference being ascribable to time spent in the laboratory. They too would like to spend less time listening and more time discussing.

Table 2

Median Allocation of Waking Hours Over Three Major Categories

a. Arts and Sciences: Male

Category		Freshmen	Sophomores	Juniors	Seniors
	N	50	19	13	11
Nonacademic	Actual	50.00	46.25	55.00	52.50
	Ideal	43.00	41.25	53.75	46.25
Academic	Actual	45.00	47.58	40.80	44.87
	Ideal	55.00	53.12	48.75	52.50
Paid employment	Actual	0	0	7.50	0
	Ideal	0	2.50	1.25	0
b. Arts and Sciences: Female					
	N	38	27	26	16
Nonacademic	Actual	53.30	46.79	50.00	45.00
	Ideal	45.50	42.62	45.50	43.55
Academic	Actual	49.15	55.71	43.50	48.30
	Ideal	55.60	55.63	57.50	55.00
Paid employment	Actual	0	0	0	0
	Ideal	0	0	0	0
c. Engineers					
	N	53	12	10	10
Nonacademic	Actual	34.15	45.00	35.00	35.00
	Ideal	32.50	40.00	31.60	35.00
Academic	Actual	58.95	51.60	61.60	53.30
	Ideal	62.00	53.30	65.00	55.00
Paid employment	Actual	0	0	0	0
	Ideal	0	0	5.00	0

Table 3

Median Allocation of Hours in Class

Category		Freshmen	Sophomores	Juniors	Seniors
a. Arts and Sciences: Male					
	N	50	19	13	11
Time spent in class	Actual	19.89	17.25	16.50	16.87
	Ideal	20.00	17.34	16.68	14.63
Time spent listening	Actual	10.33	10.16	10.25	10.75
	Ideal	9.92	9.12	9.50	9.24
Time spent discussing	Actual	3.70	2.80	3.25	.87
	Ideal	5.82	4.82	4.40	5.50
Time spent in laboratory and field work	Actual	3.82	1.45	.87	0
	Ideal	4.04	1.17	1.12	3.50
Time spent taking tests	Actual	2.04	1.94	1.70	.57
	Ideal	2.24	1.68	1.90	1.04
b. Arts and Sciences: Female					
	N	38	27	26	16
Time spent in class	Actual	19.32	18.75	16.98	19.20
	Ideal	19.05	17.86	18.00	17.25
Time spent listening	Actual	10.32	9.43	11.66	11.50
	Ideal	9.24	8.80	9.00	9.66
Time spent discussing	Actual	4.54	4.70	4.00	.85
	Ideal	6.00	5.64	6.00	4.32
Time spent in laboratory and field work	Actual	4.10	2.12	0	4.00
	Ideal	4.00	.75	.49	3.66
Time spent taking tests	Actual	2.34	2.18	.78	2.44
	Ideal	2.16	.87	.82	2.00
c. Engineers					
	N	53	12	10	10
Time spent in class	Actual	26.25	20.25	21.00	22.50
	Ideal	25.56	19.26	19.80	20.13
Time spent listening	Actual	10.24	12.00	10.60	11.00
	Ideal	8.00	9.00	11.50	9.00
Time spent discussing	Actual	5.06	3.00	4.33	1.50
	Ideal	5.86	5.66	6.00	3.00
Time spent in laboratory and field work	Actual	10.50	5.00	5.00	9.00
	Ideal	9.60	4.60	5.00	6.00
Time spent taking tests	Actual	1.44	2.25	2.00	1.50
	Ideal	1.00	2.32	1.84	1.44

Table 4 shows the breakdown of time spent outside the classroom on academic work. Medians for subgroups range from 18 to 36 hours. In no case is the median for time spent out of class as much as double the median for time spent in class. This result is notable for its failure to support the widespread belief among faculty members that students spend 2 or 3 hours studying for every hour in class.

Table 5 is concerned with the allocation of time reported by the students at the most detailed level of description. As can be seen in Table 1, at this level the description is heavily dependent on the students' subjective estimate of the qualitative character of the work. It was therefore decided to examine the data only in terms of discrepancies between actual and ideal allocations of time. In each branch (e.g., in class-listening to instructor), it was possible to distinguish between activities that could be described as self-directed and activities that could be described as teacher-directed (e.g., thinking actively and critically about the material being presented versus trying to follow the instructor). If the median self-directed activity in such a pair exceeds the teacher-directed activity, a plus sign has been entered in Table 5, and a minus sign if the reverse is true, a zero indicating no difference. In Table 5 the content of each pair has been indicated with a key word and a letter; the full text of the items to which these letters refer can be found in Table 1. The subtraction, self-directed minus teacher-directed, has been carried out separately for actual and ideal allocations of time.

Inspection of Table 5 shows that in general, the signs in the rows for actual allocations of time are negative, while the signs in rows for ideal allocations are positive. This trend is summed up in Table 6.

These results indicate that the students feel that most of their time is spent in teacher-directed activities and that they would like a shift toward more self-directed activities. With respect to activities in the classroom there is one major exception to this trend: the students would not like a shift toward tests in which major emphasis is placed on their own thought. With respect to academic work outside the classroom there is also one major exception: the students would not like a shift from the present overwhelming emphasis on assigned reading in textbooks to a broader program of unassigned reading.

Discussion

The results of this examination of how students spend their time suggest very strongly that, in the ordinary course of events, the occasion for self-directed study does not often arise at this University. This general conclusion gains support at every level of description. First, the students do not find college a very strong intellectual challenge is evidenced in their spending about as much time on nonacademic pursuits as on academic work. Secondly, the students do not spend much more time on their studies outside of class than they spend in the classroom. Thirdly, the time spent in the classroom is spent largely in listening to the instructor rather than in active discussion or laboratory work. Fourthly, the process of listening is, for these students, relatively passive and uncritical rather than active and

Table 4

Median Allocation of Hours Out of Class Spent on Academic Work

a. Arts and Sciences: Male

Category	N	Freshmen 50	Sophomores 19	Juniors 13	Seniors 11
Time spent out of class	Actual	26.13	30.36	18.50	25.12
	Ideal	34.50	34.50	28.50	31.50
Time spent reading	Actual	10.10	12.50	8.00	11.50
	Ideal	13.66	13.50	11.50	11.50
Time spent discussing	Actual	3.66	5.20	3.32	1.92
	Ideal	5.26	5.50	5.16	8.00
Time spent writing	Actual	4.52	4.24	1.19	4.25
	Ideal	5.60	4.00	2.84	5.32
Perfecting skills	Actual	5.40	6.66	1.45	1.24
	Ideal	6.40	6.75	4.25	4.00
Research and creative work	Actual	1.60	1.12	0	.16
	Ideal	3.40	5.75	1.36	.93

b. Arts and Sciences: Female

	N	38	27	26	16
Time spent out of class	Actual	28.00	32.50	28.00	30.75
	Ideal	34.50	39.15	36.00	34.50
Time spent reading	Actual	12.14	15.20	15.66	15.00
	Ideal	13.00	16.00	19.80	14.00
Time spent discussing	Actual	5.24	4.50	3.25	4.00
	Ideal	7.50	5.42	5.50	6.50
Time spent writing	Actual	4.00	.97	2.80	6.00
	Ideal	4.50	3.42	4.14	7.00
Perfecting skills	Actual	4.50	5.14	3.28	4.20
	Ideal	5.66	4.66	3.32	2.07
Research and creative work	Actual	.37	1.10	1.40	.50
	Ideal	4.04	3.16	3.50	5.00

c. Engineers

	N	53	12	10	10
Time spent out of class	Actual	31.92	27.00	36.00	30.00
	Ideal	39.36	33.00	40.00	33.00
Time spent reading	Actual	6.90	10.32	10.00	6.60
	Ideal	8.90	10.00	10.50	8.33

(Table 4 continued on next page)

Table 4 (continued)

c. Engineers (continued)

		Freshmen	Sophomores	Juniors	Seniors
Category	N	53	12	10	10
Time spent discussing	Actual	3.90	4.60	5.00	2.00
	Ideal	5.52	6.20	6.32	5.00
Time spent writing	Actual	4.56	3.00	2.00	5.00
	Ideal	4.92	.80	4.32	5.00
Perfecting skills	Actual	10.84	11.00	10.00	11.00
	Ideal	10.10	8.33	10.00	8.00
Research and creative work	Actual	.84	.30	3.00	0
	Ideal	5.30	4.00	6.50	5.80

Table 5

Preponderance of Ti[illegible]pent in Self-Directed versus Teacher-Directed Activities at t[illegible]st Subjective Level of Students' Reports

			A & S Male				A & S Female				Engineer						
Ite[illegible]			Fr	So	Jr	Sr	Fr	So	Jr	Sr	Fr	So	Jr	Sr	-	0	+
In Class																	
Listening to instructor	b - a	actual	-	-	-	-	-	-	-	-	-	-	-	-	12		
		ideal	+	+	-	+	+	+	+	+	+	-	+	+	2		10
Discussion and recitation	d - c	actual	+	-	-	+	-	-	+	-	-	-	+	+	7		5
		ideal	+	+	+	+	+	+	+	+	+	+	+	+			12
Laboratory and field work	(f+g) - e	actual	-	-	-	0	+	-	0	0	+	-	+	+	5	3	4
		ideal	+	+	+	+	+	+	+	+	+	+	+	+			12
Taking tests	j - (h+i)	actual	-	0	0	-	-	+	+	-	-	-	+	-	7	2	3
		ideal	0	+	+	-	-	+	+	-	+	-	-	-	6	1	5
Out of Class																	
Reading	(l+m) - k	actual	-	-	-	-	-	-	-	-	-	-	-	-	12		
		ideal	-	-	-	+	-	-	-	-	-	-	+	-	10		2

(Table 5 continued on next page)

Table 5 (continued)

		A & S Male				A & S Female				Engineer						
		Fr	So	Jr	Sr	Fr	So	Jr	Sr	Fr	So	Jr	Sr	-	0	+
Out of Class (continued)																
Discussing	actual	-	-	-	-	-	-	-	-	-	-	+	0	10	1	1
	o - n															
	ideal	-	+	+	+	-	+	-	+	-	0	0	0	4	3	5
Writing	actual	-	-	-	-	-	+	0	-	-	-	-	-	10	1	1
	q - p															
	ideal	+	+	-	+	+	+	+	+	-	-	+	+	3		9

Table 6

Summary of Signs of Differences: Self-directed minus Teacher-Directed Allocation of Time

	Sign of Difference		
	-	0	+
Actual	63	7	14
Ideal	25	4	55

critical. What discussion there is tends to be heavily controlled by the teacher, intended to show that the students have done their assignments, rather than a vigorous intellectual exchange. What laboratory work there is tends to be heavily supervised by the teacher, rather than an opportunity for genuinely independent work.

Outside the classroom, the qualitative picture is similar: reading, discussing, and writing are all primarily ways of carrying out specific assignments rather than opportunities for intellectual exploration and engagement. Many professors claim that this great emphasis on teacher-directed methods of education is necessary for Freshmen because they are too immature to be granted much intellectual independence. Unfortunately, the picture does not change very much from the freshman year to the senior year. For example, male Freshmen in Arts and Sciences spend a median 10.1 hours per week reading and male Seniors spend 11.5 hours per week. For both the Freshmen and the Seniors about 80 per cent of this time is spent on assignments in textbooks, leaving only 2 or at most 3 hours per week for all other kinds of reading combined. If one wonders why college graduates do not read much, an obvious answer is not far to seek. At best, the university does little to develop a taste for good reading; at worst, four years of ponderous textbooks may well produce an antipathy to all books.

The reader will have noted that many points taken up in this chapter are comparable to points considered in previous chapters. Where such comparisons are possible, results obtained from the How Students Spend Their Time Inventory are in very good agreement with those obtained from the College Characteristics Index.

In the present study, the central issue is the extent to which the experiments in self-directed study were a departure from the kind of experience to which the student in the university is usually exposed. It can be seen from the results presented in this chapter that ordinarily the student is given very little responsibility for his own education and that some of the experiments were indeed a considerable departure from this norm.

Chapter 22

The Goals of Education

The purpose of the research described in this chapter is two-fold. First, it is an effort to examine the goals of education as viewed by the students who participated in the Self-Directed Study Program. Secondly, it is an effort to detect any persistent effects of self-directed study on the participants' view of the goals of education. As reported in Chapter 18, one of the more consistent results of the 1958-59 studies was a change in attitude toward doing independent intellectual work. As compared with the controls, the experimental students recommended that a hypothetical student spend more time on activities involving independent intellectual work and less on passive-receptive learning. If this difference in attitude were to persist beyond graduation from the University, it was felt that it might be reflected in different emphases being placed on different goals of education.

The questionnaire: The Goals of Education

In developing this questionnaire on the goals of education, two problems were faced. The set of goals to be included had to be selected from a large number of possibilities, and a method had to be devised for attaching quantitative weights to these goals.

The problem of selecting goals to be listed in the questionnaire was first approached through a search of the literature. Many works were consulted, including Bloom (1956), Bryson, Finkelstein, and MacIver (1950), the report of the Educational Policies Commission (1946), and Jacob (1956), as well as the writings of Whitehead, Dewey, and other philosophers. The most immediate point of departure for the present effort was the questionnaires described in What College Students Think (Goldsen, Rosenberg, Williams, and Suchman, 1960). Those instruments could not be used intact for the present purposes, however, since they did not sufficiently single out various aspects of self-directed study for special attention. It was essential that the present questionnaire should measure the weight placed on various aspects of the goal of equipping students for independent intellectual activity after leaving college.

After much deliberation, it was decided to distinguish between two quite different types of educational goals: those referring to preparation for certain social roles (vocational, family, citizenship, etc.), and those referring to the development of certain attainments, intellectual and otherwise (knowledge, ability to think, curiosity, etc.). For the present study, the chief focus is upon the differing weights placed upon these attainments, though if preparation for social roles were not mentioned at all in the questionnaire, it was felt that respondents might view the whole instrument as artificially limited.

Accordingly, two lists of goals were prepared. The first was a list of six social roles for which college might prepare the individual: careers or occupations; family relationships; leisure time, cultural and recreational activities; religious beliefs or activities; participation as a citizen in the affairs of the community; participation in activities directed toward national or international betterment.The second list, a group of ten attainments which might be achieved through college education, is shown in Table 1. As the reader can see, goals referring to various aspects of independent intellectual activity are heavily represented in this list. This was done intentionally, since the chief purpose of this particular instrument was to determine the value placed by graduates on goals connected with self-directed study. In the actual questionnaire, the goals were presented in a haphazard order; the order in which they are shown in Table 1 is based on the results, in a manner to be explained below. As it happens, the first five goals in Table 1 are those which were intended, on an _a priori_ basis, to reflect various aspects of independent intellectual activity, while the last five are a miscellany of other goals.

The problem of attaching quantitative weights to goals was approached by means of a variation of the _constant sum method_ (see Guilford, 1954). A method similar to the one used in this study, but with a very different list of goals, was used earlier in describing nine programs for training science teachers (Gruber, 1961).

In the present study, for each list of goals, the respondents were asked to divide 100 points among the items in the list in such a manner as to indicate the distribution of emphasis among the items. An example was given involving a very uneven distribution of emphasis, in the hope that this would discourage the respondents from making very flat distributions. For each list, the respondent was asked to perform the distribution of 100 points twice: once for the emphasis placed on the goals in his own college education, and once for the emphasis that he felt should be placed on the goals in an ideal education.

Since distributing 100 points among six items is somewhat easier than distributing 100 points among ten items, the list of six social roles was presented first. This gave the respondent some practice in the use of the constant sum method; it also served to allay any concern he might have that the questionnaire was not concerned with the social roles included in the six item list. He was then asked to distribute 100 points among the items in the list of ten attainments, first to describe the distribution of emphasis on goals in his own education, then to describe the goals of an ideal college education. As it happens, of course, the present research is concerned only with the second list, and only the results for that list of goals will be reported.

Subjects. Every student who had participated in the Self-Directed Study Program, as a member of either the control or experimental groups, and had subsequently been graduated from the University, was considered eligible for inclusion. Names and correct addresses for 268 individuals were located, and in late April, 1961, each of these individuals was sent a covering letter and questionnaire. If after four weeks no reply had been received, a second letter and a second copy of the questionnaire were sent out. This procedure resulted in 157 useable questionnaires, or a response rate of 59 per cent. Of these, 92 were individuals who had been in experimental groups, and 65 in control groups. If a respondent had been in both an experimental and a control group, he was counted as an experimental subject for the present purpose, on the assumption that the control group in which he had served had had no particular effect on the character of his college education.

Results

The results are shown in Table 1, together with the exact wording of the ten goals under consideration. The goals are not listed in the order in which they appeared in the questionnaire. Instead, they are listed in the rank order of discrepancies between ideal and actual emphases, for all respondents combined. The goals at the top of the list are those on which the respondents would like to see the largest increase in emphasis, and the goals at the bottom of the list are those on which the respondents would like to see the largest decrease in emphasis.

This way of arranging the goals gives a graphic picture of the kinds of change in higher education these recent college graduates would like to see. In every case, goals which they think should receive greater emphasis have to do with self-directed engagement in intellectual activities, while the goals which they think should receive less emphasis have to do with a miscellany of activities, most of which can safely be regarded as passive or other-directed.

This pattern of discrepancies between actual and ideal was highly significant. Significance was tested with Wilcoxon's method for paired replicates: the discrepancies for the first five items in Table 1, for former controls and former experimental subjects were considered as 10 independent, signed comparisons (the second five items were omitted, because in the constant sum method, all observations are not independent). Using this approach, $N = 10$, $T = 1$, $p < .005$.

None of the differences between respondents who had been in experimental groups and those who had been in control groups approached statistical significance. Likewise, the total emphasis placed on the first 5 goals listed in Table 1 was almost identical for the two groups, for the descriptions of both their actual and their ideal educations. When these goals are arranged in rank orders for emphasis received, the rank-order correlation between experimental and control alumni responses is .90 for education actually received, and 1 for ideal education.

It should be mentioned that the variability of the results shown in Table 1 seems unduly large. This is most probably because many respondents tended to use round numbers in filling out the questionnaire. If some means of overcoming this tendency could be found, a more sensitive instrument

Table 1

The Goals of Higher Education as Seen by Alumni of the Self-Directed Study Program

Goal	Experimental N = 92 Emphasis Received				Control N = 65 Emphasis Received			
	Your own education		Ideal college education		Your own education		Ideal college education	
Increased emphasis desired	M	SD	M	SD	M	SD	M	SD
1. To develop the ability to be creative and original.	5.89	4.60	9.54	5.33	6.72	5.61	9.31	6.09
2. To develop the ability to think analytically and solve problems.	10.08	7.21	13.52	7.08	10.88	6.42	13.23	6.22
3. To develop the ability to continue to educate oneself after finishing college.	8.37	5.86	9.95	5.35	8.29	5.09	9.78	6.61
4. To establish fundamental interests which result in continuous intellectual curiosity and activity.	10.70	4.84	10.33	5.64	10.09	7.97	12.18	7.25
5. To develop the ability to form and express judgments about controversial matters and exert leadership when necessary.	7.76	5.60	8.58	4.69	7.77	5.75	8.63	4.56
Decreased emphasis desired								
6. To develop the ability to listen intelligently and profit from the knowledge and advice of experts.	8.64	4.10	8.10	4.28	8.89	4.81	7.85	3.73
7. To provide a background of knowledge of certain major concepts and ideas.	15.60	8.67	13.54	6.57	16.86	9.00	13.58	7.49
8. To develop the ability to get along with different kinds of people.	7.45	6.49	6.18	3.69	9.92	9.22	7.62	5.95
9. To develop the ability to work under supervision and carry out assignments.	7.12	5.46	5.36	3.99	7.51	5.50	5.08	3.79
10. To provide specific information and develop specific skills.	18.35	15.49	12.53	13.50	14.35	13.21	12.92	9.32

would result. In spite of this difficulty, the use of the constant sum method seems to have given a serviceable and interpretable picture of the goals of education as seen by this group.

Discussion

In spite of the consistency of the results described above, it should be mentioned that these recent graduates are, in effect, recommending only a modest change. The total discrepancy between actual and ideal education for the first five items in Table 1 (goals associated with independent intellectual activity) is less than ten per cent. Thus, the respondents would like to see about a ten per cent decrease in the sphere of passive and other-directed experiences, and a like increase in pursuit of more active, independent goals. This conservatism is reflected in the rank-order correlation between emphasis on goals in the actual and ideal education, with rho = .77 for the former experimental subjects and rho = .73 for the former controls. This group of graduates is proposing a small increase in emphasis on certain goals and a small decrease in emphasis on others. The total emphasis given the first five goals in Table 1 is just over 40 per cent for the respondents' description of their actual college education, and just over 50 per cent for their ideal of higher education.

As for the persistence of the change in attitude toward independent intellectual work found at the end of the self-directed study courses, the results are entirely negative. At the time this questionnaire was administered, two years after the experimental courses had ended, no difference between experimental and control groups could be detected. Although this is disappointing it is not surprising when one considers that each student's participation in the experimental program represented only a small fraction of his total university experience.

On the whole, wherever comparisons are possible, the results of this examination of graduates' views of the goals of higher education are in very good agreement with results presented in the two preceding chapters. In all three approaches to describing the college environment, the picture emerges of a set of educational practices and traditions which make of the student a passive recipient of what is known to his teachers without doing much to prepare him for a lifetime of creative self-education. Wherever the student is given an opportunity to criticize his education, he asks cautiously for a little more opportunity to develop his creative powers and his intellectual independence. These results gain credibility from the fact that they are very similar to results obtained with a very different group of students--high school teachers who had returned to universities for a year of additional training (Gruber, 1961).

What of the argument that the characteristics of the institution are determined by the students who come to it? In some measure, this may indeed be so. But Pace (1961) has recently cited new evidence that spending four years in an institution molds the student in the direction of the dominant environmental press; moreover, he goes on to point out that "colleges have different kinds of students because they have different kinds of programs and features" (1961, p. 12). Presumably this means that the faculty and administration are not helpless. They can alter the pattern of selective enrollment by altering the public image of the University; they can alter its public image by altering its actual program and practices.

Chapter 23

Discussion and Conclusions

This chapter is divided into two parts: a summary and discussion of the major findings, and a few recommendations. Additional discussion of numerous points is to be found in appropriate chapters throughout the report. A brief discussion of certain methodological considerations arising from the research is presented in Appendix B.

Summary of Major Findings

Performance in courses. Altogether, over a three year period, 19 courses were studied. Twelve of them involved direct comparisons between two educational treatments. In the other seven our main information is the students' and teachers' impressions, presumably based on comparisons with other educational experiences. The following series of brief sketches summarizes the entire set of results. In the left margin, next to each course listed, a plus, equal, or minus sign indicates the authors' overall impression as to whether the effort to increase the student's responsibility for his own education had a favorable, negligible, or negative effect on the student's learning of the subject matter of that course. Not included in this evaluation are the students' own reactions to the course, or changes in attitude toward doing independent intellectual work: these matters will be taken up separately.

Summary of Nineteen Courses, 1957-1959

+ Chapter 2, Educational Psychology

Experimental group slightly superior on final and on 15 month posttest of retention of subject matter. On posttest experimental group showed more curiosity about educational psychology.

+ Chapter 3, General Psychology

No basis for comparison on final examination. On one-year posttest experimental group retained more, showed greater curiosity about subject.

+ Chapter 4, Civil Engineering: Reinforced Concrete Design

No comparison group available, instructor delighted with projects.

+ Chapter 5, Freshman English

Group receiving no classroom training in grammar superior on final test of grammar.

+ Chapter 6, Introduction to Research Methods in Psychology

No difference in final examination, tendency for students exposed to self-directed study in first part of course to continue with second part of course.

+ Chapter 7, Physical Optics

Experimental group gave superior pattern of performance on final examination and on three-month posttest.

+ Chapter 8, Culture and Personality
No comparison group available. Instructor felt that group reports were excellent, suggested that self-directed study coupled with careful instructor guidance is very successful.

= Chapter 8, General Psychology
No comparison group available. Instructor felt that experimental group learned about as much as it would have with conventional treatment.

+ Chapter 8, Independent Study in Honors
No final examinations. Students and faculty very pleased.

= Chapter 9, Freshman English
Groups about equal on final examination.

- Chapter 10, Physical Science
Experimental group slightly inferior on final examination.

- Chapter 11, Principles of Economics
Experimental group markedly inferior in final examination.

- Chapter 12, Principles of Marketing
Experimental group slightly inferior on final examination.

= Chapter 13, Political Science: American Government
Groups about equal on final examination.

+ Chapter 14, Principles of Chemical Engineering
Experimental group definitely superior on final examination and all other tests.

+ Chapter 15, General Psychology - I
Control group definitely superior on final examination. Experimental group equal to Control group on one year posttest of retention of factual knowledge, superior on test of critical thinking about psychology.

- Chapter 16, General Psychology - II
Control group definitely superior on final examination.

+ Chapter 17, Mathematics: Analytical Geometry and Calculus
No comparison group available. Instructor felt that experimental group had lost in proficiency in certain skills, gained in depth; he felt method was good for superior students.

+ Chapter 17, Advanced Chemical Engineering Thermodynamics
No comparison group available. Instructor felt that self-directed study equipped students very well for still more advanced courses.

The results for 1958-1959 should be singled out for special attention, since it was in that year that the attempt was made to conduct many experiments according to a single plan and to test the pattern hypothesis, i.e., the hypothesis that experimental students would do their best work on the parts of the final examination which required them to solve complex problems and learn new material, while control students would do their best work on parts of the final which involved factual knowledge and simple problems (see Part III, p. 1).

The results for seven courses conducted in the Fall of 1958 are shown in Table 1. These courses are Freshman English (see Chapter 9), Physical Science (see Chapter 10), Principles of Economics (see Chapter 11), Principles of Marketing (see Chapter 12), Political Science (see Chapter 13), General Psychology - I (see Chapter 15), and General Psychology - II (see Chapter 16). The results are treated in standard score form in order to make comparisons possible among diverse courses and tests. Table 1 shows the mean standard score for each part of the final and for the total final examination for all experimental students and for all control students. The columns headed "percentile rank" translate the standard scores into percentile form (assuming a normal distribution) to give a clearer indication of the magnitude of the differences between the groups. For example, in Part I of the final, the mean experimental student was at the 44th percentile (for the two groups combined), while the mean control student was at the 56th percentile.

Table 1

Mean Standard Scores for Final Examination Performance

of Seven Experimental Courses in 1958-1959

		Part I		Part II		Part III		Part IV		Total	
	N	Mean	Percentile rank	Mean	Percentile rank	Mean	Percentile rank	Mean	Percentile rank	Mean	Percentile rank
Experimental	534	-.16	44	-.20	42	-.29	39	-.09	46	-.22	41
Control	594	.15	56	.16	56	.26	60	.05	52	.18	57

It can be seen from Table 1 that the Control group was superior to the Experimental group for these seven courses. In the total final examination performance the magnitude of the difference was sufficient to put the mean experimental student at the 41st percentile and the mean control student at the 57th percentile. While this difference is not necessarily catastrophic, it is very distinct. Using a method proposed by Wilkinson (1951) for dealing with a situation where a number of tests of significance is carried out, the superiority of the control group is significant in five out of these seven courses; in the other two (Freshman English and Political Science) the differences between groups are negligible.

Table 1 also indicates that the pattern hypothesis was not supported. There is some tendency for the experimental and control groups to be most alike on Part IV of the final examinations, but they are most dissimilar on Part III, rather than on Parts I and II as required by the hypothesis. In the limited sense that the experimental subjects did their best work on Part IV (learning and applying new material), the hypothesis is not actually contradicted. But genuine support for the hypothesis would require that the experimental subjects equal or surpass the controls on Part IV.

The reader will notice that quite different impressions are conveyed by the summary of all 19 courses on the one hand and Table 1 on the other. In the former, 12 courses are rated favorable to self-directed study, 3 as neutral, and 4 as unfavorable. The data on which Table 1 are based involve 5 courses unfavorable to self-directed study, and 2 neutral. What accounts for the difference? A number of points must be taken up.

1. Table 1 omits the course in Principles of Chemical Engineering (Chapter 14), because in that course only a two-part final examination was used. There is, however, little doubt as to the superiority of the self-directed study group over the control group.

2. Table 1 is concerned only with final examination performance. For General Psychology - I (Chapter 15), the experimental group was considered superior to the control because, in spite of the former's inferiority on the final examination, it was superior on the one year posttest.

3. Even in courses of an introductory nature comparable to most of those considered in Table 1, experiments conducted in the earlier years tended to produce results more favorable to self-directed study. It may well be that the sacrifice of control of the treatment variable implicit in extending the breadth of the experimental program was responsible for this effect. For further details about this loss of control see discussions in Chapters 9, 11, and 16.

4. In the summary, five of the six courses without a comparison group available received a "plus" rating, and the sixth received an "equal" rating. It is conceivable, of course, that these favorable impressions are unduly subjective. An alternative interpretation is that courses without control groups tended to be more advanced courses (since there is often only one section taught in an advanced course it is harder to arrange controls), and that self-directed study actually had a more favorable outcome in more advanced courses. This idea is borne out by the favorable results for Physical Optics and for Principles of Chemical Engineering, the two most advanced courses in which control groups were available.

It seems that the more advanced the course, the more successful is SDS. But it should be noted that a course may only be "advanced" in the sense that the students in it have encountered earlier versions of the same material. Juniors being introducted to a new subject, such as economics, do badly in SDS; Freshmen continuing with a subject previously studied, such as English, do well in SDS. Results for Physical Optics and Chemical Engineering bear this out. Both of these courses cover material reminiscent of earlier courses, and self-directed study worked well in both.

Why should this be? Perhaps it is because the student has picked up a collection of models of thought and study from his previous experience in the subject. Those serve to carry him along in the subject in question: he has internalized them enough so that he can be independent within a subject that he has been shown how to study. But the models he has developed are not sufficiently general to permit him to tackle a new field independently.

It might require only minor modifications of present teaching methods to get from such specialized models to more general ones: just a little attention by each teacher to the training of his students in independent work might pay off very quickly.

Attitudes and Self-Directed Study. In this section we will consider first the effects of self-directed study upon certain attitudes, and then the influence of the student's attitudes upon his performance in self-directed study. Results concerning certain personality variables will also be treated in this section.

In a very real sense, the finding that self-directed study increased the student's curiosity about the subject studied, as measured by the rise in questioning behavior, should be thought of as a change in attitude. In this case, the change in attitude is, as far as the evidence goes, a restricted one, limited to the particular subject matter of the experimental course. Increased questioning behavior was found in the experimental group both in the student's thought-processes during the lecture itself (Chapter 19) and in posttests long after the experimental course was over (Chapters 2 and 3). In addition, there was an increased tendency for students who were required to do independent work to go on to a further course in the same area (Chapter 6). It would be interesting to know whether the increase in curiosity resulting from self-directed study generalized to other subjects.

A more general change in attitude was the increase in emphasis placed on the value of doing independent intellectual work by students who had participated in experimental classes. This finding is all the more striking when it is recognized that this change toward a more favorable attitude toward academic independence occurred in a group of seven courses in which the students' evaluations of the experimental courses were unfavorable, and in which their final examination performance was inferior to the control group.

A generally negative reaction to participation in the self-directed study experiments was expressed in the students' evaluations of the courses. In almost every case, students prefer the conventional method. Morale problems associated with similar experiments have been reported by Eglash (1954), Neel (1959), and Beach (1961). One student, quoted by Eglash, said, "This method won't work unless we are brought up in this system and are used to it, and unless everyone cooperates. It allows too much independent thinking" (1954, p. 261). There seems to be a uniformly negative initial reaction to methods of instruction giving the student greater independence than he is accustomed to. In some cases, as reported by Beach and by Eglash, this first reaction gives way to a more favorable one during the experimental course. In other cases, as reported by Neel, by Baskin & Churchill (1961), and in the present study, the student's reaction remains somewhat unfavorable even at the end of the course. But his unfavorable evaluation of the particular self-directed study course in which the student has been involved must be interpreted in light of the development of a more positive general attitude toward independent intellectual work.

Perhaps the sequence of events can be epitomized, if not explained, as follows: (a) the initial exposure to changed teaching methods is unpleasant simply because it is a change away from accustomed and comfortable methods; (b) nevertheless, the widening of the student's range of experience with teaching methods leads him to take a more accepting attitude toward independent work; (c) but there is a kind of "emotional lag" in which the unfavorable initial reaction to the particular course persists even though the generalized attitudes underlying it have already changed as a result of exposure to the experiment; (d) only when the student has clearly adopted a new set of expectations regarding proper educational methods can he be expected to gain the maximum benefit from self-directed study.

A more general theoretical formulation of the process which may be occurring during sequences such as those described above has been offered by Festinger (1957), who argues that the individual will respond to "dissonance" between his opinions and actions he is required to carry out, by changing his opinions in such as way as to reduce the dissonance. This formulation was tested and substantiated by Festinger and Carlsmith (1959), who found, in addition, that the maximum opinion change occurred when the pressure used to elicit the required overt behavior was just sufficient to do so; amounts of pressure greater than this necessary minimum produced smaller changes in opinion. Thus, the rate of movement through the sequence of attitudinal and behavioral changes involved in adjusting to new educational methods may depend on subtle variations in the teacher's behavior, on the way in which the method in question deviates from institutional norms, and on other factors determining the felt amount of pressure toward behavior change.

With regard to the effect of personality variables on the student's response to self-directed study, no clear or simple picture has emerged. The most commonly advanced hypothesis, that students with a high need for independence will prefer and do better in situations affording them more opportunity for independent work, received little or no support in the work of Koenig and McKeachie (1959). In the present study, this hypothesis was examined in some detail in Chapter 7, with very meager results. It may not be possible to predict directly from broad personality variables to performance in particular situations. In a somewhat more promising approach, Vroom (1959) found what amounts to a triple interaction between personality, performance, and satisfaction. Thus, given a measure of the individual's need for independence and a measure of the extent of his participation in a decision-making process, it was possible to predict something about his satisfaction. While interactions of this kind are interesting, they would provide no basis for deciding how to assign the individual to an educational treatment.

That we are very far from a clear understanding of the relation between student personality and teaching methods is suggested by a surprising result recently reported by Amidon and Flanders. In an eighth grade group, they found that dependent-prone children learned more from methods involving less direction by the teacher than from a more directive method, while independent-prone students were not particularly affected by this variable of teaching method. The authors suggest the following interpretation: "As the teacher becomes more directive, this type of student finds increased satisfaction in more compliance, often with less understanding of the problem-solving steps carried out. Only when he is free to express his doubts, to ask questions and gain reassurance, does his understanding keep pace with his compliance to the authority figure. Lacking this opportunity, compliance alone may become a satisfactory goal and content understanding may be subordinated to the process of adjusting to teacher directives" (Amidon and Flanders, 1961, p. 290). Nothing in the present authors' experience suggests that this comment is _in_applicable to college students.

If we restrict our attention to the relation between fairly specific aspects of the individual's ideology and his response to self-directed study, a somewhat clearer picture emerges. In the study mentioned above, Koenig and McKeachie found that "...students who think the teacher should be authoritarian tended to do poorly in independent study" (1959, p. 134). Bending and Hountras found that "The authoritarian would be happier in a well-structured course where the structure is imposed by the department" (1959, p. 6). In the present study, although the correlations involved were very low, there was some tendency for students having a higher regard for independent intellectual work to do better work and to give more favorable course evaluations if they were in self-directed study groups than if they were in conventional groups (see Chapter 18).

If we consider intellectual independence as an aspect of the growth of the individual, the question arises, when is the individual ready for self-directed study? Crutchfield (1962) found that college sophomores were the most yielding, or conforming, of all college students, second only to military officers among the groups he studied. In Chapters 15 and 18 above, it was shown that in one course, so far as final examination performance was concerned, freshmen suffered more than sophomores from self-directed study, but so far as changes in attitude toward independent intellectual work was concerned, freshmen gained from exposure to the experimental method, while sophomores did not. Thus, if we assume that the preferred sequence is: change in values leads to spontaneous change in behavior, it may be necessary to initiate the sequence by an enforced change in behavior. Self-directed study may be best introducted in the freshman year, not because it leads at once to better performance but because it leads to the most rapid change in educational values. As Bruner has summed up this way of thinking, "Passion, like taste, grows on its expression. You more likely do yourself into feeling rather than feel yourself into doing" (1962).

Academic ability and self-directed study. One of the most commonly advanced hypotheses proposes that the abler or brighter student may be able to profit from methods of teaching giving him greater responsibility for his own education, while the less able student requires closer supervision and guidance. In their study of undergraduates at Antioch, Baskin and Churchill found no support for this hypothesis. They found "no evidence that the independent study methods needed to be reserved for the superior or advanced student only" (1961, p. 4). The present study provided a setting in which to examine the same question with a much wider assortment of students.

On the whole, the results of the present study provide little or no support for the notion of a direct relation between intellectual ability and capacity to profit from self-directed study. In the group of seven courses included in the correlational analysis reported in Chapter 18, the correlation between intelligence and final examination performance was almost identical for experimental and control groups. As reported in Chapter 15, in a course in General Psychology there was an interaction between ability and performance on one part of the final such that the upper 20 per cent of students on ability seem to have profited from exposure to self-directed study. The particular form of the results gave no reason to believe that students of lower ability suffered from exposure to self-directed study. In the course in Physical Optics (Chapter 7), there was some evidence that students of lower general ability profited the most from exposure to self-directed study, especially if they had a high academic aspiration level. Students of higher ability did not do as well in the experimental group as in the control group. There was reason to believe that some students of high ability have learned how to succeed in college without really trying, and that when these individuals are confronted with a departure from accustomed procedures they are particularly handicapped. This result is similar to an earlier finding that students of very high ability did not do as well in an advanced teacher training program as students of somewhat lower ability (Gruber, Brady, and Means, 1959). All in all, these results suggest that interrelationships among intellectual ability, difficulty level of the material being studied, educational method, and various attitudinal factors may be too complex to provide any basis for assigning individual students to educational treatments. To these complexities should be added the recurrent finding in recent studies that there may be little relationship between creativity and ability as measured by typical intelligence tests (see, for example, Holland, 1961; Torrance and Harmon, 1961). Even if it is deemed wise to assign superior students to special programs, it may not be easy to decide upon appropriate criteria for superiority.

In view of the foregoing, it may be suggested that we replace the hypothesis of a direct relationship between measures of general intelligence and ability to profit from independent intellectual work with the following two hypotheses: Hypothesis 1 - Students who are accustomed to intellectual struggle will perform best when such struggle is required of them. Thus if a change in teaching methods is introduced, not the most facile but the most determined students will display the highest level of performance. But high performance is not identical with maximum intellectual growth. This consideration leads to Hypothesis 2 - In a situation where students are not stimulated to peak intellectual performance, those students will profit most from a renewed challenge who have been most somnolent in the past, provided the challenge is definite and protracted enough to ensure the appropriate readjustments on the students' part. From this point of view, educators face two questions simultaneously. What conditions of learning foster the most rapid acquisition of knowledge? What conditions of learning foster the most positive response to new intellectual challenges? If the two questions receive different answers, a third question arises: what mixture of methods produces the optimal balance between the acquisition of knowledge and the positive response to new challenges? Or is there some method, yet to be invented, that eliminates the conflict between these distinct educational goals?

Recommendations

The picture of the institution that emerges from the present study is of an environment fundamentally hostile to independent intellectual work on the student's part. Therefore, no single innovation in education method could be expected to produce the profound change in atmosphere necessary to give the student greatly enlarged responsibility for his own education nor could a single change adequately nurture his ability to persevere in a lifetime of intellectual independence.

Therefore, it should come as no surprise that limited changes in educational method produce only limited and evanescent effects. The present study, for example, focusses attention on the effects of reducing the number of lectures per course from about three per week to one per week, seemingly a radical change. But if we look at this change from the student's point of view it is not so very great. He is still enmeshed in the course system; the total number of formal classes he attends per week has been reduced from about 15 to about 13;he is still attending an institution characterized by predominantly teacher-directed academic work; the experiment in which he participates lasts for only one semester. Much bolder variations in educational method than those involved in the present study can readily be conceived. If the researcher wishes to study large changes in intellectual development he may have to discover or invent variations in institutional practices based on the sort of pervasive change that only occurs when a whole institution is given over to the pursuit of a new set of goals.

The following suggestions are intended to give some indication of how such a pervasive change might be introduced. Some of them follow quite directly from the present research, some only indirectly. They have been selected from a large number of possibilities with two aims in view: first, to indicate how appropriate changes might be made in the methods of higher education at a variety of points in a complex system; secondly, to indicate that some useful changes might be made quite easily while others would involve more radical revisions of current practices.

Orientation to college. The goal of developing intellectual independence should be clearly stated to the students as one of the prime purposes of higher education. This could be done in many forms, in printed matter sent to prospective students and in other orientation programs. A carefully worked out summer reading program could be developed, to get the new student started on independent intellectual work before he arrives at college. This might be implemented by practical means of making recommended books available to entering students. The point of such a program would not be simply to squeeze in a little extra reading, but to provide a concrete embodiment of an important educational objective, a form of action by which the student could begin promptly to move toward that objective.

Reading courses. Courses of study in which the student receives academic credit for carrying out a purposeful reading program without attending classes connected with the subject in question now exist, but they could be expanded to become an important part of the university scene. It would be possible to require every student to do some of his work in this form.

Dismissal of lectures. In courses typically taken by freshmen and sophomores, classes could be dismissed for a week or two, and students required to learn at least a portion of each subject without close supervision by the instructor. A rotation system could readily be developed to avoid dismissing classes in all fields at the same time. With adequately prepared students, more radical reductions in the number of formal class meetings could be smoothly accomplished in junior and senior courses.

Reduction in number of courses. The prevalent system in which the student takes five or six courses per semester presents a real obstacle to the development of intellectual independence. From the student's point of view, when his work is fragmented in so many pieces he is probably well-advised to be compliant, to observe carefully what the instructor wants of him and do no more than that. If he takes an enthusiastic plunge into something that interests him he may forget to apportion his study time equally among his many subjects. Detached compliance is a safer posture. Reducing the number of courses a student takes would permit him to combine respectful attention to his instructors' wisdom with some pursuit of knowledge in a more highly individualized way.

Toward planfulness. No matter how capable the individual is of working without supervision, he cannot be accurately called "independent" unless he has a well developed and well used capacity for forming his own purposes and plans. Academic programs now give virtually no attention to developing this capacity. Although some effort in this direction would seem to be essential, the present authors are somewhat baffled as to what techniques to propose.

If, by methods such as those outlined above, a concerted effort were made to use the first two years of college to prepare the student to assume greater responsibility for his own education, the last two years would probably be marked by the spontaneous emergence of a great variety of techniques embodying the same ideal. In sum, the problem is not so much to disseminate any single educational technique as to elevate the goal of developing intellectual independence to a central place where it can play a governing role in every decision a teacher makes.

References

Aikin, W. M. (ed.) *Thirty schools tell their story*. New York: Harper, 1943.

Amidon, E. and Flanders, N. A. The effects of direct and indirect teacher influence on dependent-prone students learning geometry. *J. educ. Psychol.*, 1961, 52, 286-291.

Astin, A. W. A re-examination of college productivity. *J. educ. Psychol.*, 1961, 52, 173-178.

Attneave, F. *Applications of information theory to psychology*. New York: Holt-Dryden, 1959.

Barker, R. G., & Wright, H. F. *Midwest and its children: the psychological ecology of an American town*. New York: Row, Peterson, 1955.

Baskin, S. and Churchill, Ruth. Experiment in independent study (1956-1960). *Antioch College Reports*, No. 2, 1961.

Beach, L. R. The use of instructorless small-groups in a social psychology course. Paper read at Western Psychological Association, Seattle, June, 1961.

Bending, A. W. and Hountras, P. T. Anxiety, authoritarianism, and student attitude toward departmental control of college instruction. *J. educ. Psychol.*, 1959, 50, 1-7.

Berlyne, D. E. A theory of human curiosity. *Brit. J. Psychol.*, 1954, 45, 180-191. (a)

Berlyne, D. E. An experimental study of human curiosity. *Brit. J. Psychol.*, 1954, 45, 256-265.

Berlyne, D. E. *Conflict, arousal and curiosity*. New York: McGraw-Hill, 1960.

Birney, R. C. and Taylor, M. J. Scholastic behavior and orientation to college. *J. educ. Psychol.*, 1959, 50, 266-274.

Bloom, B. S. Thought processes in lectures and discussions. *J. gen. Educ.*, 1953, 7, 160-169.

Bloom, B. S. (ed.) Taxonomy of educational objectives: the classification of educational goals; handbook 1: cognitive domain. New York: Longmans, Green. 1956.

Bonthius, R. H., Davis, F. J., and Drushal, J. G. The independent study program in the United States. New York: Columbia Univer. Press, 1957.

Brinkley, S. G. Mental activity in college classes. J. educ. Res., 1953, 46, 535-541.

Bruner, J. S. The conditions of creativity. In Gruber, H. E., Terrell, G., and Wertheimer, M. (eds.) Contemporary approaches to creative thinking. New York: Atherton Press, 1962 (in press).

Bryson, L., Finkelstein, L., and MacIver, R. M. (eds.) Goals for American Education. New York: Harper, 1950.

Crutchfield, R. S. Conformity and creative thinking. In Gruber, H. E., Terrell, G. and Wertheimer, M. (eds.) Contemporary approaches to creative thinking. New York: Atherton Press, 1962 (in press).

Dressel, P. L. and Mayhew, L. B. General education: explorations in evaluation. Washington, D. C.: American Council on Education, 1954.

Duncan, D. B. A significance test for differences between ranked treatments in an analysis of variance. Virginia J. of Science, 1941, 2, 171-189.

Educational Policies Commission. Policies for education in American democracy. Washington, D. C.: National Education Association of the United States, 1946.

Edwards, A. L. Experimental design in psychological research. New York: Rinehart, 1950.

Edwards, A. L. Edwards personal preference schedule, revised manual. New York: The Psychological Corporation, 1959.

Eglash, A. A group-discussion method of teaching psychology. J. educ. Psychol., 1954, 45, 257-267.

Eisner, S. and Rohde, K. Note taking during or after the lecture. J. educ. Psychol., 1959, 50, 301-304.

Festinger, L. A theory of cognitive dissonance. Evanston, Ill.: Row, Peterson, 1957.

Festinger, L. and Carlsmith, J. M. Cognitive consequences of forced compliance. J. abnorm. soc. Psychol., 1959, 58, 203-210.

Goldsen, Rose K., Rosenberg, M., Williams, R. M., and Suchman, E. A. What college students think. Princeton: Van Nostrand, 1960.

Gruber, H. E., Brady, K. P., and Means, J. R. Toward the improvement of science teaching. Univer. Colorado Behav. Res. Lab. Rept. No. 9, 1959.

Gruber, H. E. Science as thought: a study of nine Academic Year Institutes for the training of science teachers. Univer. Colorado Behav. Res. Lab. Rept. No. 16, 1961.

Gruber, H. E. The Farrand Hall experiment. Univer. Colorado Behav. Res. Lab. Rept. No. 17, 1961.

Gruber, H. E. and Weitman, M. Item analysis and the measurement of change. J. educ. Res., 1962, in press.

Guilford, J. P. Psychometric methods. (2nd ed.) New York: McGraw-Hill, 1954.

Holland, J. L. The prediction of college grades from the California Psychological Inventory and Scholastic Aptitude Test. J. educ. Psychol., 1959, 50, 135-142.

Holland, J. L. Creative and academic performance among talented adolescents. J. educ. Psychol., 1961, 52, 136-147.

Hovey, D. E. The development of curiosity under an experimental method of teaching the educational psychology course. Unpublished M.A. thesis, Univer. of Colorado, 1958.

Jacob, P. E. Changing values in college. New Haven: Hazen Foundation, 1956.

Karsten, A. Psychische Sättigung. Psychol. Forsch., 1928, 10, 142-254. Summarized in Lewin, K. Dynamic Theory of Personality, New York: McGraw-Hill, 1935.

Katona, G. Organizing and memorizing: studies in the psychology of learning and teaching. New York: Columbia Univer. Press, 1940.

Kersh, B. Y. The adequacy of "meaning" as an explanation for the superiority of learning by independent discovery. J. educ. Psychol., 1958, 49, 282-292.

Koenig, Kathryn and McKeachie, W. J. Personality and independent study. J. educ. Psychol., 1959, 50, 132-134.

Lindquist, E. F. Statistical analysis in educational research. Boston: Houghton Mifflin, 1940.

Liverant, S. The use of Rotter's Social Learning Theory in developing a personality inventory. Psychol. Monogr., 1958, 72, No. 2, 1-23.

McConnell, T. R. and Heist, P. The diverse college student population. In Sanford, N. S. (ed.) The American College, New York: Wiley, 1962.

McKeachie, W. J. Student-centered versus instructor-centered instruction. J. educ. Psychol., 1954, 45, 143-150.

McKeachie, W. J. The improvement of instruction. Rev. educ. Res., 1960, 30, 351-360.

McKeachie, W. J. Procedures and techniques of teaching: a survey of experimental studies. In Sanford, N. S. (ed.) The American College, New York: Wiley, 1962.

Mearns, Hughes. Creative power: The education of youth in the creative arts. 2nd revised edition, New York: Dover, 1958.

Milton, G. A. and Lipetz, M. E. The factor structure of human needs as measured by the EPPS. Unpublished paper, 1961.

Murray, H. A. Explorations in personality. New York: Oxford Univer. Press, 1938.

Nachman, M. and Opochinsky, S. The effects of different teaching methods: a methodological study. J. educ. Psychol., 1958, 49, 245-249.

Neel, Ann F. The relationship of authoritarian personality to learning: F Scale scores compared to classroom performance. J. educ. Psychol., 1959, 50, 195-199.

Pace, C. R. Psychological differences between college environments. College Board Review, 1960, 41, 24-28.

Pace, C. R. The validity of the College Characteristics Index as a measure of college atmosphere. Paper presented at American Psychological Association, New York, 1961.

Pace, C. R. and Stern, G. G. An approach to the measurement of psychological characteristics of college environments. J. educ. Psychol., 1958, 49, 269-277.

Pace, C. R. and Stern, G. G. College Characteristics Index. Dittoed supplement to 1958 preliminary manual. 1960.

Stern, G. G. Preliminary manual: College Characteristics Index. Psychol. Res. Center, Syracuse Univer., 1958.

Stern, G. G. Environments for learning. In Sanford, N. S. (ed.) The American College, New York: Wiley, 1962.

Stroud, J. B. Experiments on learning in school situations. Psychol. Bull., 1940, 37, 777-807.

Thistlethwaite, D. L. College press and student achievement. J. educ. Psychol., 1959, 50, 183-191.

Thorndike, E. L. The teaching of controversial subjects. Cambridge: Harvard Univer. Press, 1937.

Torrance, E. P. and Harmon, J. A. Effects of memory, evaluative, and creative reading sets on test performance. J. educ. Psychol., 1961, 52, 207-214.

Vroom, V. H. Some personality determinants of the effects of participation. J. abnorm. soc. Psychol., 1959, 59, 322-327.

Walker, Helen, & Lev, J. Statistical inference. New York: Holt, 1953.

Wilkinson, B. A statistical consideration in psychological research. Psychol. Bull., 1951, 5, 156-158.

Appendix A

General Scheme for Final Examinations

I. Information (knowledge of subject matter)
Questions asking what, when, where, who, and requiring brief, factual answers.

Examples: Who wrote The Possessed?
What is Boyle's Law?
Where did the Populist movement originate?
When did the Industrial Revolution begin?
What is a loss-leader?
Who promulgated the doctrine of laissez-faire?
When does heat exchange occur?
Where is the optic chiasma?

II. Simple Applications
Tasks requiring utilization of one general principle and knowledge of a few facts.

Examples: A steel ball is falling toward the earth at a speed of 50 m.p.h. when at an altitude of 1000 feet. What is its speed at an altitude of 800 feet?

A small, partly industrialized, capitalist country is in the throes of a depression. This country has always been plagued by overpopulation and chronic unemployment. What would Lord Keynes recommend?

III. More Difficult Applications
Tasks requiring utilization of two or more principles and fairly broad factual knowledge.

Examples: In the past election Byron Johnson mildly surprised the experts by defeating John Mackie for the office of U. S. Representative. How did this come about?

You want to build a pendulum clock which will be an extremely accurate timepiece. Briefly describe the important problems you will encounter and your solutions to these problems.

A nationwide chain of shoe stores is considering moving into the Boulder area. Where ought they locate this store? Explain.

IV. Learning and Application of New Material

New material is introduced as part of the final examination. This material should be related to the rest of the subject matter covered in the course but never before dealt with in the course. The students' task is to read and then apply this new material to the solution of a problem. Solving this problem should require integration of the new information with the rest of the material and skills learned in the course.

Appendix B

Methodological Notes

The following paragraphs are all concerned with difficulties arising from the attempt to conduct controlled classroom experiments within a single institution. Some of these questions have recently been discussed by McKeachie (1962).

The placebo effect. Probably the most frequent question asked in connection with studies of the kind described in the present report refers to the "Hawthorne effect," that is, the possibility that improvements in performance stemming from the experimental method are due simply to the fact that the participants are aware of the fact that they have been singled out for special treatment. This awareness supposedly leads to an improvement in interest and morale which in turn has a beneficial effect upon performance. In so far as results can be ascribed to the placebo effect, the specific character of the experimental treatment does not matter, so long as the participants in the experiment are aware of receiving differential treatment. There is reason to believe that the placebo effect cannot explain the present results, since in many cases a deterioration in morale was coupled with an improvement in performance. Nevertheless, the point is inescapable that experiments within a single institution are complicated by the participants' awareness of differential treatment, and the more radical the experiment, the more thoroughly aware of it they become.

Assignment of teachers to treatments. Using the same teacher in experimental and control treatments is not as satisfactory as might appear. The teacher may not be able to keep two different roles distinct; if there are more than two treatments the problem is further complicated. In any case, a teacher can have only one philosophy of education at a time, and the effects of asking him to teach in a way that violates his own views are not well understood. On the other hand, using different teachers for different treatments is not a satisfactory solution. If the teachers are permitted to volunteer for different treatments, there is a volunteer error; if the teachers are assigned at random to treatments the relation between the teacher's own views and the treatment variable becomes uncontrolled and unpredictable. Moreover, random assignment of teachers to treatments is not very feasible, since many teachers are dedicated individuals with strong beliefs about methods of teaching. Resistance generated by attempts to employ arbitrary procedures artificially limits the scope of many experiments. One solution to this problem is always to arrange matters so that the teacher can teach in accordance with his own views; this means that experimental variations can only be introduced by changing the teacher's views. This task is not at all insuperable but it may be far easier to accomplish by changing the orientation of an entire institution than by singling out a few individuals for special attention.

The measurement of change. As Nachman and Opochinsky (1958) have pointed out, the effect of an experimental variation may be concealed by the insensitivity of the technique of measurement. Sensitivity can be increased by testing promptly, rather than after long delays, and by devising more sensitive tests. The question remains, do we really want sensitive devices for picking up small effects of cautious experiments? If the point is to develop new methods of teaching that might actually be worth adopting, perhaps it is better to use long-range tests and crude instruments, placing our reliance on maximizing the independent variable, i.e., introducing radical changes with more readily detectable effects. Again, such large variations are not easy to introduce within a single institutional framework.

On the need for experimental field studies. From all these considerations it follows that the problems of control in research involving large variations in educational method may not be solvable within a single institution. It therefore becomes necessary to consider methods of research involving more than one institution. One approach to this problem, of course, is through quantitative field studies of intact institutions. It may be possible to increase the power of such field studies by introducing systematic experimental variations, once the field setting is well enough understood.